PROTEINS: Structure, Function, and Genetics 47:236–249 (2002)

Comparative Structural Analysis of Psychrophilic and Meso- and Thermophilic Enzymes

Giulio Gianese,[1] Francesco Bossa,[1] and Stefano Pascarella[1,2]*

[1]*Dipartimento di Scienze Biochimiche, A. Rossi Fanelli, and Centro di Biologia Molecolare del C.N.R., Università La Sapienza, Roma, Italy*

[2]*Centro Interdipartimentale di Ricerca per l'Analisi dei Modelli e dell'Informazione nei Sistemi Biomedici (CISB), Università La Sapienza, Roma, Italy*

ABSTRACT **Enzymes adapted to cold display structures comparable with those of their meso- and thermophilic homologs but are characterized by a higher catalytic efficiency at low temperatures and by thermolability at moderate temperatures. To identify the structural factors responsible of such features, we undertook a systematic comparative analysis of several structural properties in a data set consisting of 7 cold active enzymes belonging to different structural families and 28 related structures from meso/thermophiles representing most of the structural information now available. Only high-resolution and high-quality structures were considered. Properties were calculated and then compared for each pair of 3D structures displaying different temperatures of adaptation using a temperature-weighting scheme. The significance of the resulting differences was evaluated with a statistical method. Results reveal that each protein family adopts different structural strategies to adapt to low temperatures. However, some common trends are observed: the number of ion pairs, the side-chain contribution to the exposed surface, and the apolar fraction of the buried surface show a consistent decrease with decreasing optimal temperatures. Proteins 2002;47:236–249.**

Key words: structural differences; cold adaptation; protein stability; electrostatic interactions; surface properties; solvation free energy; secondary structural variations; packing density; protein cavities

INTRODUCTION

The earth's surface presents a wide spectrum of different environments whose colonization by life required a large variety of adaptive strategies. Enzymes synthesized by organisms living in habitats characterized by extreme physicochemical conditions are of great interest for the high adaptive pressure they underwent. The most widespread extreme condition for life is represented by low temperatures. Indeed, the biosphere is dominated by environments that, from an anthropocentric point of view, are considered cold, 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, these enzymes display improved catalytic efficiency at low temperatures reflected in their higher turnover number (rise in k_{cat}) and/or increased affinity for the substrate (reduction of K_M).

It is generally assumed that the structural determinants of such kinetic features are responsible not only for the low temperature optima but also for the inactivation at moderate temperatures (usually >40°C) of the cold active enzymes. It has been proposed that the thermolability of cold enzymes is a consequence of a higher flexibility of their molecular structure, in contrast with the thermostability of enzymes from thermophiles that conversely should be correlated with rigidity of their polypeptide chain. Because of their higher catalytic efficiency at low temperatures, enzymes extracted from psychrophilic organisms were investigated for their biotechnological potential, in particular in industrial processes as energy savers and in detergent industry as additives.[4,5]

Many theoretical and experimental studies have been devoted to clarify the molecular basis of the adaptation of thermophilic enzymes to high temperatures, comparing single thermostable proteins with their mesophilic homologs and systematically examining different properties for families of proteins. Interesting results have been achieved through comparative analysis.[6,7] On the other hand, molecular mechanisms of cold adaptation are still relatively unknown. Until recently, the absence of available crystal structures of cold enzymes limited the comparison to their homologs from meso- and thermophilic organisms only to comparative modelling and multiple sequence alignments.[8–10] In earlier work, we performed a comparative sequence analysis to detect common and general

Grant sponsors: Ministero dell'Istruzione, dell'Università e della Ricerca.

*Correspondence to: Stefano Pascarella, Dipartimento di Scienze Biochimiche, Università La Sapienza, P.le A. Moro, 5, 00185 Roma, Italy. E-mail: Stefano.Pascarella@uniroma1.it

Received 6 July 2001; Accepted 30 November 2001

strategies of adaptation to low temperatures.[11] Using 15 homology-modeled structures and 6 crystal structures of psychrophilic enzymes along with multiple alignments of their sequences with homologs from meso- and thermophilic species, we detected preferred amino acid substitutions and variations in amino acid residue composition possibly related to adaptive strategies to low temperatures. The observed exchanges were also correlated to different structural environments.

The recent and growing amount of available structural and sequence data on cold active enzymes prompts a detailed structural comparative analysis. Crystal structures have been solved for the cold enzymes: α-amylase from *Altheromonas haloplanctis* A23,[12] citrate synthase from an Antarctic bacterium, strain DS2-3R,[13] elastase and trypsin from Atlantic salmon,[14,15] glyceraldehyde-3-phosphate dehydrogenase from lobster,[16,17] malate dehydrogenase from *Aquaspirillium arcticum,*[18] triosephosphate isomerase from *Vibrio marinus*,[19] and an alkaline protease from an Antarctic *Pseudomonas aeruginosa,* strain TACII18.[20] Single-family examinations were already performed by several authors, through comparison of these structures with their meso- and thermophilic counterparts and identification within each protein family of the trends involving structural features possibly related to cold adaptation.

We report a systematic comparative analysis among the crystallographic structures of psychrophilic enzymes and their meso- and thermophilic homologs. The data set consists of 7 cold adapted and 28 meso- and thermophilic protein structures distributed in 7 structural families. Each structure in the collection was examined for several structural properties. Results of calculations were compared for each pair of three-dimensional structures displaying temperature differences of their source organisms. Changes in structural features were measured with a temperature-weighting scheme and their significance was evaluated by using a statistical scheme.

MATERIALS AND METHODS

Collection of Data Sets

The crystallographic structures of the available cold active enzymes were collected from the Protein Data Bank (PDB).[21] Homologous structures from meso- and thermophylic organisms were retrieved from PDB with the program FASTA.[22] Only unique structures were considered. In the presence of alternative structures for the same protein, only that displaying the best resolution, possibly without point mutations, and passing the PROCHECK[23] and WHAT_CHECK[24] quality assessments, was considered. Sequences of the selected proteins were then structurally aligned to each psycrophylic homolog with the program CE.[25] The alignments were manually corrected by inspection of the superimposed structures. To ensure structural homology, only sequences sharing ≥30% residue identity to each other were included. A total of 104 incomplete side-chains distributed in 5 proteins (Table I) were rebuilt by the program HOMOLOGY in the package InsightII.[26] The side-chain rotamer displaying the lower nonbond energy was kept and treated as experimental. Ligands (cofactors, inhibitors, substrate analogs, etc.) and solvent molecules were always removed from the structures. Optimum growth temperature was assigned to each protein by considering the normal living environmental temperature (or the average of a range of normal habitat temperatures) of the corresponding species, in the case of monocellular and ectothermic organisms. Homeothermic organisms were set at 37°C. Use of host environmental temperatures is significant and has several precedents in literature.[6,27]

Evaluation of the Properties

The trends versus growth temperature of several chemicophysical properties were calculated by pairwise comparison of the proteins in each family. Each property difference was divided by the difference between the associated temperatures of the two proteins so as to obtain the change per temperature unit. This change was then weighted by the square of the same temperature difference. The sum of the weighted differences per Celsius degree derived from all the possible pairwise comparisons within the family was finally normalized by dividing for the sum of the square of all the temperature differences considered. In a given family, the score S for a certain property per 10°C fall is so determined by the equation proposed by Vogt et al.[6]:

$$S = \frac{\sum_{i=2}^{n}\sum_{j=1}^{i-1}(T_i - T_j)^2\left[\frac{X_i - X_j}{|T_i - T_j|}\right]10}{\sum_{i=2}^{n}\sum_{j=1}^{i-1}(T_i - T_j)^2} \quad (1)$$

X_i and X_j are the values of the property for the proteins i and j of the family, respectively; T_i and T_j are the normal living temperatures of the host species; n is the number of members in the family. Elements with the same temperature do not contribute to the sum because the corresponding temperature difference is equal to zero. The assignment of the family components to a position i is effected by following the decreasing temperature order: $i = 1$ corresponds to the molecule with the warmest temperature and $i = n$ to the molecule with the coldest temperature (psychrophile). Consequently, in Eq. 1, T_i will be necessarily lower or equal to T_j. S measures the variation of the property X by 10°C fall toward cold temperatures; thus, $S < 0$ denotes decreasing X and vice versa.

The final score S_M is the mean over all the family S. To assess the statistical significance of S_M, the assigned optimal temperatures were randomly shuffled[6] in each family, and S_M was recalculated for each property. The shuffling was repeated 1000 times, and the final mean and standard deviation of the distribution of the randomized scores S_M were calculated. The scores can be expressed in terms of Z-score:

$$Z_X = \frac{S_M - \bar{S}_r}{\sigma_r} \quad (2)$$

TABLE I. List of Enzymes Used in the Work

Family name	Source organism	Growth temperature (°C)	PDB ID	Structure resolution (Å)	Identity[a] (%)	No. of subunits	Sequence length (monomer)	No. of rebuilt side-chains
1. α-amylase	*Alteromonas haloplanctis* A23	4	1aqm	1.85	—	1	448	1
	Tenebrio molitor	25	1jae	1.65	45	1	471	—
	Porcine (pancreatic)	37	1dhk	1.85	43	1	496	—
	Human (pancreatic)	37	1hny	1.80	42	1	496	—
	Human (salivary)	37	1smd	1.60	42	1	496	—
2. Citrate synthase	Antarctic bacterium DS2-3R	5	1a59	2.09	—	2	377	20
	Pyrococcus furiosus	110	1aj8	1.90	40	2	371	—
3. Elastase	*Salmo salar*	4	1elt	1.61	—	1	236	—
	Human (leukocyte)	37	1ppf	1.80	38	1	218	—
	Porcine (pancreatic)	37	1qnj	1.10	68	1	240	—
4. Glyceraldehyde-3-phosphate dehydrogenase	*Homarus americanus*	20	4gpd	2.80	—	4	333	—
	Leishmania mexicana	37	1a7k	2.80	58	4	358	—
	Escherichia coli	37	1gad	1.80	65	4	330	—
	Bacillus stearothermophilus	55	1gd1	1.80	53	4	334	—
	Thermus aquaticus	72	1cer	2.50	49	4	331	60
	Thermotoga maritima	82	1hdg	2.50	50	4	332	—
5. Malate dehydrogenase	*Aquaspirillium arcticum*	4	1b8p	1.90	—	2	327	—
	Porcine (heart)	37	5mdh	2.40	51	2	333	—
	Thermus flavus	72	1bmd	1.90	62	2	327	—
6. Triosephosphate isomerase	*Vibrio marinus*	15	1aw2	2.65	—	2	255	—
	Saccharomices cerevisiae	27	7tim	1.90	43	2	247	—
	Human	37	1hti	2.80	41	2	248	—
	Chicken	37	1tph	1.80	41	2	245	—
	Trypanosoma cruzi	37	1tcd	1.83	41	2	248	—
	Escherichia coli	37	1tre	2.60	65	2	255	—
	Plasmodium falciparum	37	1ydv	2.20	38	2	246	—
	Leishmania mexicana	37	1amk	1.83	39	2	250	—
	Trypanosoma brucei	41	1tpf	1.80	39	2	250	—
	Bacillus stearothermophilus	55	2btm	2.40	42	2	250	22
	Thermotoga maritima	82	1b9b	2.85	39	4	252	—
7. Trypsin	*Salmo salar*	4	2tbs	1.80	—	1	222	—
	Porcine	37	1fni	1.60	66	1	223	—
	Rat	37	3tgi	1.80	67	1	223	1
	Human	37	1trn	2.20	64	1	224	—
	Bovine	37	1c1t	1.37	66	1	223	—

[a]Percentage of residue identity to the cold active homolog.

where $\bar{S}$ and σ_r are the randomized mean and standard deviation for the score S_M of the property X.

For electrostatic (hydrogen bonds, ion pairs, charge at helix ends) and geometric (overall packing, buried residue packing, cavity volume, buried water volume) properties, only subunits were considered to avoid the contribution of the subunit-subunit interactions derived from the families with oligomeric proteins, whereas for the remaining properties, protein structures corresponding to the biological unit were considered.

All the programs were written in C-language or Perl and run under the IRIX 6.3 operating system.

Hydrogen Bonds and Ion Pairs

Hydrogen bonds were calculated by using the program HBPLUS[28] with the default parameters, except for the maximum donor-acceptor distance, set to 3.5 Å instead of 3.9 Å, to be closer to that proposed by Baker and Hubbard[29] (3.1–3.2 Å). Atoms of opposite charge separated by a distance < 4 Å and not involved in H-bond were assigned to an ion pair. This distance threshold is generally accepted after systematic analysis of a large sample of protein structures.[30] Atoms with positive charge included side-chain nitrogens in Arg and Lys; atoms with negative charge were the side-chain oxygens of Asp and Glu. Ion pairs involving His were not considered because of the ambiguous assignment of its protonation state in the proteins. Complex salt bridge, defined as ion-paired interaction joining more than two side-chains,[31] and simple salt link involving two side-chains but more than two charged atoms were not considered in the calculation as a single interaction, but rather each individual atomic interaction (single ion pair) was counted. For example, if one side-chain oxygen

of an Asp is connected to both η-nitrogens of an Arg, we count a total of two ion pair interactions, and if one of the η-nitrogens and the ζ-nitrogen of the same Arg are interacting to one side-chain oxygen of an another Asp, then we count a total of four interactions.

Because the families in our data set display different quaternary structures, H-bonds and ion pairs were counted only in isolated subunits. However, the calculations for ion pairs were performed on all the other available subunits of a protein to detect possible interactions missing in the first monomer. The maximum number of ion pairs observed was taken as the final count for the monomer. The final counts were normalized by the number of residues in the subunit. Ion pairs involving charged oxygens or nitrogens at chain ends were not considered because one or more amino acids in these regions were missing in a few structures.

Surfaces, Hydrophobic Contacts, and Solvation Free Energy

Accessible surfaces for all, apolar, and side-chain atoms were calculated by using the program NACCESS[32] with the default probe radius of 1.4 Å.

Atomic buried surfaces were calculated by the software package OS (Occluded Surface) that implements the occluded surface method.[33] According to this procedure, an atomic dot surface corresponding to the van der Waals surface is generated with the algorithm described by Connolly.[34] At each dot surface is associated a distance normal. The normal is extended until it intercepts the van der Waals surface of an adjacent atom: the dot surface associated to the normal is then considered occluded by the intercepted atom. However, surface dots bearing normals longer than 2.8 Å are by definition nonoccluded. The total occluded surface of an atom is defined as the sum of the elementary occluded dot surfaces. Apolar buried surface was considered as the sum of the occluded surface of carbon and sulfur atoms, whereas apolar contact surface was defined as the portion of apolar buried surface occluded by carbon and sulfur atoms.

The solvation free energy of protein folding (ΔG_s) was calculated by using the atomic solvation parametric (ASP) approach. Such a method assumes that the effect of the solvent is proportional to the solvent accessible surface area buried on protein folding. The solvent contribution can be expressed in terms of the solvent accessible surface area multiplied by the solvation free energy per unit area as defined through the ASP set. The accessible surface was examined through the individual surface contribution of each atom type. A solvent energy factor (cal mol^{-1} $Å^{-2}$) was assigned to each atom type. This factor measures the energy variation when a certain atom is transferred from an apolar to a polar solvent. The total solvation energy results from the sum of the products between the exposed surface areas of the single atoms and the corresponding energy factors. Several authors have proposed different ASP sets for each atom type. A systematic comparison of 9 different ASP sets[35] suggested the most reliable be that proposed by Eisenberg et al.[36] named "sch1." The program ASC (Analytic Surface Calculation)[37,38] with the default parameters (1.4 Å probe radius) was used for the calculation of the exposed surfaces for each atom type. The solvation free energy of the folded state (ΔG_f) relative to the transfer of the molecule from an apolar to a polar solvent was computed as the sum of the products of atomic accessible surface areas given by ASC and the associated solvation factors of the "sch1" set. The solvation free energy of the unfolded state (ΔG_u) was calculated by summing the transfer energy of each residue. Such an energy was defined for each residue through the application of ASC with the "sch1" ASP set to an extended Gly-X-Gly peptide, where X is one of the 20 amino acids. The solvation free energy of the protein folding derives from the difference between the transfer energy of the folded and unfolded state: $\Delta G_s = \Delta G_f - \Delta G_u$. Finally, the free energy was normalized by the total number of residues of the protein, because of dependence on polypeptide chain size.[35]

Secondary Structural Variations

Secondary structures were assigned by the program STRIDE.[39] Residue conformations identified with the symbols H or G were considered as helices. Residues designated with E or B were considered β-strands. Remaining symbols were assigned to nonrepetitive regions. Insertions and deletions were also included, counting one variation for each residue deleted or inserted. The rate of transition from one secondary structure to another was calculated by Eq. 1. $X_i - X_j$ was replaced by the number of residues changing from one secondary structure type to another in each pairwise protein comparison within each family. The rate was averaged over all families, and the net flux of exchanges per 10°C fall in temperature was calculated. The calculation was repeated for every possible transition between conformational symbols.

Packing Density, Cavity Volume, and Buried Water

Packing density was calculated with the above mentioned program package OS[33] that uses the normal associated with each dot of a generated atomic dot surface and extended to the van der Waals surface of the close atom. The length of this distance is divided by 2.8 to obtain a normalized packing parameter *PP*, equal to zero for atoms in contact, and equal to 1.0 or more for atoms separated by at least the distance corresponding to the diameter of a water molecule. The average over the *PP* associated to all the dots of a residue gives the residue packing parameter *PPr*. The mean of the *PPr* of all the residues of a protein is the average protein packing parameter *PPp*. Packing density is defined as $1-PPp$. The same calculation was performed on the subset of residues at no more than 0.05 fractional accessible surface,[40] to determine the packing density of the protein core.

Protein volumes were calculated with the software GRASP,[41] whereas volumes, surfaces, and number of internal cavities were measured by using the server CASTP (http://cast.engr.uic.edu/cgi-bin/cast1/index.pl) developed for detection of protein pockets and cavities with a

solvent probe of 1.4 Å radius.[42] The overall cavity volume was determined by summing all the van der Waals volumes enclosed by wall atoms and without mouth openings. Total cavity volume was then normalized by the molecular volume enclosed in the protein van der Waals surface.

Potential buried water in protein interior was assigned by the suite DOWSER[43] that locates internal cavities and assess their hydrophilicity in terms of the free energy of interaction with a water molecule. Only water molecules with binding energy below -10 kcal mol^{-1} are retained by the program. Zhang and Hermans[43] found that cavities with interaction energies below -12 kcal mol^{-1} tend to be filled, whereas cavities with energies greater than that threshold tend to be empty. The total number of water molecules was then multiplied by the mean cavity volume per water in a protein structure (25.6 $Å^3$), as reported by Hubbard et al.[44] The total volume occupied by buried water was then normalized by the molecular volume.

Charge Distribution at Helix Ends

Only helices consisting of more than two turns (i.e., longer than seven residues) were considered. The geometric center of the helix N-end was calculated from the coordinates of the C_α, N, and C atoms of the main chain relating to the first three residues at the N-terminal and the N-cap. The geometric center of the C-end was determined in the same way from the last three residues at the C-terminal and the C-cap. All the charges detected within a 7 Å distance from the geometric centre were summed. Atoms with positive charge refer to the nitrogens of the Lys and Arg side-chains, with a formal charge of $+1$ and $+1/3$, respectively, for each N atom. Negative charged atoms are identified in the oxygens of the Asp and Glu side-chains, with a formal charge of $-1/2$ for each O atom. The average charge at N- or C-ends of helices was calculated by dividing the sum of the charges inside the N- or C-end spheres of all the selected helices by the number of helices. This average charge is the property analyzed in Eq. 1.

RESULTS

Data Set

The selection procedure described in Materials and Methods yielded a set of seven families containing 7 psychrophilic, 21 mesophilic, and 7 thermophilic structures accounting for a total of 35 enzymes (Table I). They share at least 38% sequence identity. All of them passed the PROCHECK[23] and WHATCHECK[24] quality controls except the psychrophilic glyceraldehide-3-phosphate dehydrogenase (GAPDH; PDB entry 4gpd) for which no suitable alternative was available. For this reason, the GAPDH family was excluded from the ion pair and H-bond calculations, which are most sensitive to the quality of the structure.

Statistical Test

The significance of the observed differences in structural properties was measured by a Z-score (see Materials and Methods), assuming a normal distribution of the randomized S_M scores. Values of Z-score ≥ 2.0 or Z-score ≤ -2.0, corresponding to a p-value of about 2%, were considered significant. The distributions of the randomized S_M scores for the structural features discussed in the work are reported in Figure 1.

Hydrogen Bonds and Ion Pairs

Statistics relative to the total and the per-residue number of hydrogen bonds and ion pairs observed in isolated subunits of our data set are reported in Table II. As discussed above, the GAPDH family was excluded from the calculations because of poor quality. Three of the six families display a decrement of the total and per-residue number of H-bonds toward cold temperatures. The average number of H-bonds lost every 10°C fall toward cold temperatures is -2.9 and -0.0009 on a per-subunit and per-residue basis, respectively. The contribution to the average score from the different types of H-bonds (i.e., those involving atoms of main-chain/main-chain, side-chain/side-chain, and main-chain/side-chain) is -0.7, -0.6 and -1.6, respectively (Table II). Total and per-residue number of ion pairs decrease in four of the five families considered (the trypsin family was not included because of lack of ion pairs). On average, -1.1 and -0.0027 ion pairs are lost every 10°C fall per subunit and per residue, respectively (Table II). The α-amylase family displays the highest drop rate in the per-subunit number of H-bonds (-17.9) and ion pairs (-3.6) toward cold temperatures.

Similar trends are observed when entire biological units are considered (data not reported).

Surfaces, Hydrophobic Contacts, and Solvation Free Energy

Statistics about the change in fractional accessible surface for different atom types (apolar and side-chain) are reported in Table III. Five of the seven families show an increasing fraction of apolar accessible surface with decreasing optimal temperature. The trend is generally rather weak except for the GAPDH family, which displays an increase of 0.0052 in fraction of apolar accessible surface per 10°C fall. Two families, α-amylase and citrate synthase, behave differently, yet only the former bears a clear decrease of the fraction of apolar surface equal to a loss of -0.0085 per 10°C fall. The average score for the whole set of families is equal to 0.0011. Because of the complementarity of polar and apolar accessible surfaces, the tendency for the former is the inverse of that of the latter. The contribution of side-chain atoms to the accessible surface decreases in six of the seven families, with an average of -0.0022 per 10°C drop in temperature. The total accessible surface decreases in four families, whereas the average score for whole data set corresponds to a -55.7 $Å^2$ loss per 10°C fall in temperature.

The analysis of the variation in the fractional apolar occluded (i.e., buried) surface is reported in Table III. Five of the seven families display a decreasing fraction of apolar occluded surface. The average loss of such fraction over all families is -0.0007 per 10°C fall.

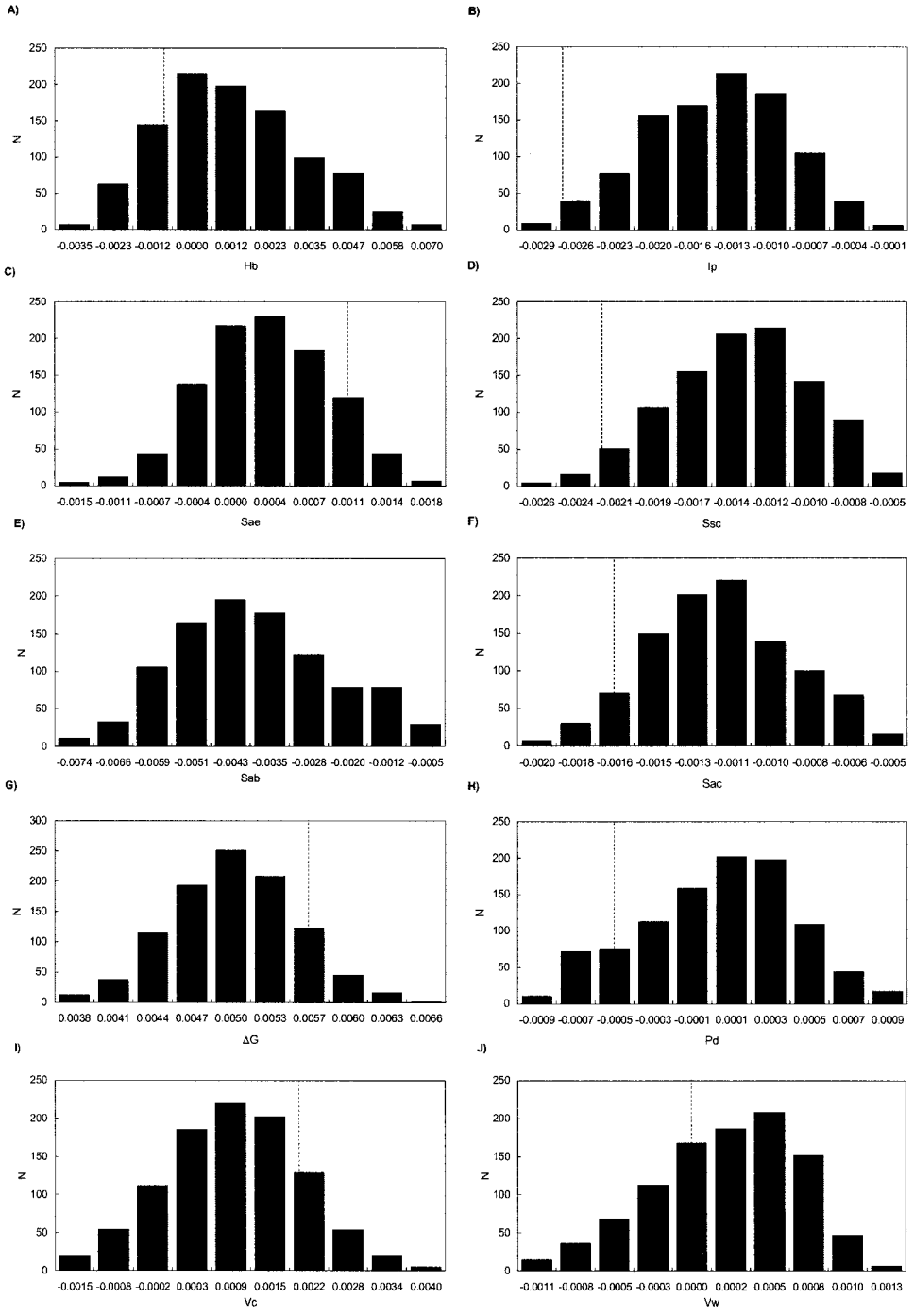

Fig. 1. Histograms of the number N of temperature randomized scores S_M falling in the S_M intervals reported on the x axis for the (**A**) per-residue number of hydrogen bonds (Hb), (**B**) per-residue number of ion pairs (Ip), (**C**) fraction of apolar exposed surface (Sae), (**D**) fraction of side-chain accessible surface (Ssc), (**E**) fraction of apolar buried surface (Sab), (**F**) fraction of apolar contact surface (Sac), (**G**) solvation free energy of folding normalized per residue (ΔG), (**H**) packing density (Pd), (**I**) fraction of cavity volume (Vc), and (**J**) fraction of buried water volume (Vw). Mean value of each interval is shown on the x axis. Values on the x axis in histograms (E), (I), and (J) are multiplied by 10. Dotted line indicates the position on the x axis of the real S_M score for the relative property.

The fractional apolar contact surface is calculated as the ratio between the apolar contact and the total occluded surfaces. It decreases in six of the seven families, whereas in the elastase family the rate is zero. The average score is −0.0016 per 10°C fall toward low temperatures.

Solvation free energy of protein folding ΔG_s increases in five of the seven families toward cold temperatures (Table III). The average rate over all families is 0.0057 kcal mol^{-1} per residue per 10°C fall in living temperature.

TABLE II. Number of Hydrogen Bonds and Ion Pair Interactions in the Data Set Proteins

Family[a]	PDB-ID	Temperature (°C)	H-bonds per subunit	M-M	S-S	M-S	H-bonds per residue	Ion pairs per subunit	Ion pairs per residue
1.	1aqm	4	436	237	74	125	0.973	20	0.044
	1jae	25	456	246	72	138	0.968	25	0.053
	1dhk	37	468	252	68	148	0.943	33	0.066
	1hny	37	493	257	83	153	0.993	28	0.056
	1smd	37	518	252	95	171	1.044	34	0.068
	Score		−17.9	−5.1	−2.7	−10.1	−0.0068	−3.6	−0.0060
2.	1a59	5	349	224	48	77	0.925	23	0.061
	1aj8	110	346	235	31	80	0.932	12	0.032
	Score		0.3	−1.0	1.6	−0.3	−0.0007	1.0	0.0028
3.	1elt	4	192	119	18	55	0.813	0	0.000
	1ppf	37	158	101	21	36	0.724	2	0.009
	1qnj	37	201	123	21	57	0.837	2	0.008
	Score		3.8	2.1	−0.9	2.6	0.0098	−0.6	−0.0026
4.[b]	—	—	—	—	—	—	—	—	
5.	1b8p	4	292	208	31	53	0.892	15	0.045
	5mdh	37	323	218	37	68	0.969	17	0.051
	1bmd	72	323	219	39	65	0.987	23	0.070
	Score		−4.5	−1.6	−1.2	−1.7	−0.0139	−1.2	−0.0037
6.	1aw2	15	223	163	24	36	0.874	18	0.070
	7tim	27	232	152	19	61	0.939	24	0.097
	1hti	37	217	156	15	46	0.875	28	0.112
	1tph	37	235	145	28	62	0.959	17	0.069
	1tcd	37	229	148	25	56	0.919	12	0.048
	1tre	37	244	171	25	48	0.956	22	0.086
	1ydv	37	234	147	22	65	0.951	27	0.109
	1amk	37	249	155	26	68	0.996	9	0.036
	1tpf	41	250	155	27	68	1.000	12	0.048
	2btm	55	240	163	19	58	0.960	17	0.068
	1b9b	82	219	149	27	43	0.858	28	0.109
	Score		0.8	0.9	−0.4	0.3	0.0051	−1.0	−0.0039
7.	2tbs	4	177	109	20	48	0.797	0	0.000
	1fni	37	181	108	19	54	0.811	0	0.000
	3tgi	37	179	109	19	51	0.802	0	0.000
	1trn	37	173	108	17	48	0.772	0	0.000
	1c1t	37	176	105	24	47	0.789	0	0.000
	Score		−0.1	0.4	0.1	−0.6	0.0011	0.0	0.0000
	Average score		−2.9	−0.7	−0.6	−1.6	−0.0009	−1.1	−0.0027
	Z-score		−1.9	−1.2	−2.0	−1.8	−1.0	−2.5	−2.2

[a]Numbers refer to the family numbers in Table I.
[b]GAPDH family was not included in calculations.

Secondary Structural Variations

The marginal decrease of main-chain/main-chain hydrogen bonds (Table II) should reflect the conservation of secondary structural content on adaptation to low temperatures. Indeed, the analysis of variation of secondary structural content per 10°C decrease in host temperature does not indicate any significant flux (Fig. 2). Only a weak trend is observed for decreasing β-strands (−0.69). Instead, deletions tend to increase toward cold temperatures (1.22).

Packing Density, Cavity Volume, and Buried Water

The change rates of packing density, fractional cavity volume, and fractional volume of buried water for all the protein families are listed in Table IV. Values of packing density range from a minimum of 0.365 (PDB entry 1ydv) to a maximum of 0.421 (PDB entry 1jae), with a large fluctuation within each family. Five of the seven families display a tendency to packing reduction with decreasing temperature. However, the average rate is only −0.0005 per 10°C fall toward cold temperatures. Packing of the buried residues is more conservative, with a rate of −0.0003. Single protein values range from a minimum of 0.498 (PDB entry 1cer) to a maximum of 0.564 (PDB entry 1a59), and only three of the seven families decrease the packing of the hydrophobic interior. There is no obvious correlation between the variation rates of the packing of

TABLE III. Change in Total and in Fraction of Total Solvent Accessible Surface Contributed by Apolar Atoms and Side-Chain Atoms*

Family[a]	PDB-ID	Temperature (°C)	Apolar accessible surface	Side-chain accessible surface	Total accessible surface (Å^2)	Apolar occluded surface	Apolar contact surface	ΔG_s (kcal mol^{-1})
1.	1aqm	4	0.522	0.811	16242.2	0.752	0.589	−0.839
	1jae	25	0.512	0.818	16061.3	0.753	0.590	−0.832
	1dhk	37	0.547	0.827	17914.1	0.760	0.598	−0.849
	1hny	37	0.541	0.829	17794.5	0.756	0.594	−0.846
	1smd	37	0.553	0.829	17641.0	0.756	0.595	−0.830
	Score		−0.0085	−0.0054	−509.0	−0.0017	−0.0021	0.0011
2.	1a59	5	0.566	0.851	26933.1	0.749	0.578	−0.858
	1aj8	110	0.568	0.870	26596.4	0.767	0.606	−0.977
	Score		−0.0002	−0.0018	32.1	−0.0017	−0.0027	0.0113
3.	1elt	4	0.523	0.819	9708.7	0.768	0.619	−0.888
	1ppf	37	0.522	0.825	10207.5	0.772	0.627	−0.916
	1qnj	37	0.500	0.831	10478.2	0.763	0.611	−0.902
	Score		0.0036	−0.0027	−192.2	0.0001	0.0000	0.0064
4.	4gpd	20	0.610	0.835	47741.9	0.753	0.446	−0.846
	1a7k	37	0.594	0.837	47518.0	0.750	0.402	−0.860
	1gad	37	0.556	0.829	45442.5	0.753	0.455	−0.856
	1gd1	55	0.531	0.832	44706.3	0.754	0.449	−0.893
	1cer	72	0.556	0.840	45539.2	0.755	0.458	−0.941
	1hdg	82	0.582	0.843	45354.8	0.752	0.453	−0.893
	Score		0.0052	−0.0014	368.8	−0.0002	−0.0035	0.0123
5.	1b8p	4	0.575	0.846	25020.6	0.756	0.590	−0.881
	5mdh	37	0.591	0.861	26029.4	0.757	0.598	−0.870
	1bmd	72	0.561	0.853	25468.5	0.754	0.597	−0.879
	Score		0.0021	−0.0010	−63.6	0.0003	−0.0010	−0.0003
6.	1aw2	15	0.568	0.827	19005.5	0.751	0.585	−0.829
	7tim	27	0.567	0.846	19017.1	0.750	0.585	−0.836
	1hti	37	0.580	0.840	18725.1	0.752	0.588	−0.816
	1tph	37	0.578	0.830	18830.2	0.750	0.589	−0.834
	1tcd	37	0.571	0.851	19101.3	0.751	0.590	−0.893
	1tre	37	0.585	0.842	19510.3	0.742	0.579	−0.816
	1ydv	37	0.513	0.864	20378.5	0.754	0.590	−0.861
	1amk	37	0.601	0.850	19230.8	0.745	0.579	−0.856
	1tpf	41	0.588	0.842	19202.2	0.748	0.587	−0.871
	2btm	55	0.549	0.854	18103.4	0.755	0.590	−0.876
	1b9b	82	0.566	0.870	19774.8	0.759	0.598	−0.894
	Score		0.0014	−0.0056	−46.7	−0.0016	−0.0021	0.0107
7.	2tbs	4	0.556	0.826	9515.2	0.764	0.612	−0.857
	1fni	37	0.537	0.810	9410.3	0.764	0.613	−0.859
	3tgi	37	0.548	0.820	9497.2	0.765	0.610	−0.859
	1trn	37	0.537	0.835	9769.8	0.765	0.609	−0.845
	1c1t	37	0.547	0.807	9110.1	0.768	0.618	−0.842
	Score		0.0042	0.0024	20.7	−0.0004	−0.0001	−0.0017
	Average score		0.0011	−0.0022	−55.7	−0.0007	−0.0016	0.0057
	Z-score		1.4	−2.0	−1.1	−2.3	−1.5	1.3

*The change in fraction of total occluded surface contributed by apolar atoms and by apolar interacting atoms; per residue variation in solvation free energy of protein folding.
[a]Numbers refer to the family numbers in Table I.

the hydrophobic core and that of the overall protein. Indeed, buried residues typically display values of packing density larger and less variable than the solvent accessible residues, having scarce influence on the overall protein packing.[45]

Fractional volume of cavities increases only in three of the seven families, with the greatest rate displayed by the GAPDH family (0.0015), even though the bad quality of the crystal structure of the psychrophilic enzyme could affect the family score. The average score over all the

		COLD				
		H+G	E+B	NR	Del	Sum
WARM	H+G		0.09	1.93	0.15	2.17
	E+B	0.11		2.06	0.37	2.55
	NR	1.85	1.65		2.10	5.61
	Ins	0.37	0.12	0.93		1.41
	Sum	2.33	1.86	4.92	2.63	
	Net sum	0.16	-0.69	-0.69	1.22	
	Z-score	-0.7	-1.5	-1.0	2.0	

Fig. 2. Number of secondary structural substitutions found in the multiple sequence alignments of psychrophilic (cold) and mesophilic or thermophilic (warm) enzymes. The transitions from one conformational state to another were determined for each family as described in Materials and Methods. Data reported represents the mean over all the family scores. Sums in the column correspond to exchanges from a structural state in a meso/thermophile to all other possible states in the psychrophile. Sums in the row correspond to variations from psychrophiles to meso/thermophiles. The net sum (row sum − column sum) corresponds to net flux toward one secondary structure from all others in the direction from warm to cold temperatures. Secondary structural conformations considered were helices [α-helix (H) + 3_{10}-helix (G)], β structures [strand (E) + β-bridge (B)] and nonrepetitive regions (NR). Insertions (Ins) and deletions (Del) were also considered.

families is only 0.0002 per 10°C fall. The family trends of fractional cavity volume are in agreement with the rates of change of the cavity surface and cavity number (data not reported).

The fractional volume of buried water does not show significant changes, except for the α-amylase (−0.0012), and the average rate per 10°C fall tends to zero. The program DOWSER correctly assigns all the water molecules experimentally observed inside the protein structures of the data set, bearing a predicted interaction energy less than −10 kcal mol^{-1}, with the only exception of 7 of a total of 454. The experimental buried water molecules represent the 67.3% of the total number of those assigned by DOWSER. The correlation between predicted and crystal water is independent of the structure resolution.

Charge Distribution at Helix Ends

The hydrogen bonds and peptide groups in α-helices point in the same direction. For this reason, a cumulative effect derives from the dipoles of each peptide bond; consequently, α-helix should have a macrodipole moment whose interactions with charged atoms or polar groups can increase protein stability.[46] To analyze the correlation between living temperature and helix stabilization, charge distribution at α-helix ends was examined. Charged atoms within a sphere of 7 Å from the N- and C-cap of the helices of each protein of our collection were calculated, and the total charge was than divided for the number of helices considered. The result reveals that charge distribution within each family does not differ significantly in enzymes adapted to different environmental temperatures. However, our analysis confirms a clear preference for a net negative charge to occur at the N-terminus, except for elastase and trypsin families, which have a low helix fraction. The charge distribution at the C-terminus is less consistent.

DISCUSSION

We undertook a systematic comparison of several physicochemical and structural properties between psychrophilic and homologous meso/thermophilic proteins to detect possible family-specific and/or general trends in the cold adaptation of enzymes. In a previous work[11] we analyzed the net flux of amino acid substitutions from meso/thermophilic to psychrophilic enzymes. This work was also aimed at a structural interpretation and rationalization of such observations.

A criticism may be raised to most comparative analyses of proteins concerning the fact that the enzyme structures considered are often from organisms found in extremely divergent phylogenetic lineages, and in many cases even from different domains of the evolutionary tree. Distantly related organisms probably were subjected to several and different selective pressures due to their individual phylogenetic history, being reflected on enzyme function and structure, which could not be related to temperature adaptation.[47] Ideally, intermediate steps during adaptation to low temperatures should be examined; unfortunately, only extant lineages can be compared. However, precautions may help balancing the potential phylogenetic bias. In our case, only structures sharing at least 38% sequence identity are included in the sample (Table I). This similarity level ensures[48] a strong conservation of the main structural features. Moreover, the weighting scheme based on ΔT and the statistical test applied should help filtering the phylogenetic noise that may exist in the data.

As reported by several authors,[6,49–52] thermophilic proteins tend, in general, to possess more electrostatic interactions than their mesophilic counterparts, indicating this structural feature as one of the major determinants of the improved thermal stability. The same conclusion is reached by Szilágyi and Závodszky[7] for ion pairs, although they criticize the hypothesis of Vogt et al.[6] that the increase of H-bonds be the most important stabilizing factor at high temperatures, because it was flawed by the inclusion of low-quality structures in their data set. However, the importance and the role of hydrogen bonds in protein stability remain still controversial.[53,54] Presumably, these interactions contribute to the rigidity of protein structure,[55] whereas their loss should increase flexibility, a feature thought to be correlated to cold adaptation.[2] The average rate of variation of H-bonds per subunit over all the families has a Z-score equal to −1.9, close to the threshold value of significance. Thus, although this parameter probably does not represent a significant stabilizing factor at high temperatures, it cannot be excluded that loss of H-bonds may play an important role in adaptation in the opposite direction (i.e., cold) of the biological temperature range. The major contribution to this rate comes from the H-bonds among side-chain atoms (S-S in Table II) that display a Z-score equal to −2.0. The reduction of the number of such interactions could weaken the restraints

TABLE IV. Change in Packing Density, in Buried Residue Packing, in Fraction of Total Cavity Volume, and in Fraction of Total Structural Water Volume for Each Family

Family[a]	PDB-ID	Temperature (°C)	Packing density	Buried residue packing	Cavity volume	Buried water volume
1.	1aqm	4	0.399	0.524	0.021	0.017
	1jae	25	0.421	0.547	0.017	0.020
	1dhk	37	0.407	0.527	0.022	0.021
	1hny	37	0.419	0.533	0.020	0.021
	1smd	37	0.419	0.536	0.021	0.021
	Score		−0.0044	−0.0018	−0.0001	−0.0012
2.	1a59	5	0.390	0.564	0.009	0.008
	1aj8	110	0.401	0.561	0.012	0.009
	Score		−0.0010	0.0003	−0.0003	−0.0001
3.	1elt	4	0.384	0.529	0.020	0.015
	1ppf	37	0.374	0.525	0.016	0.013
	1qnj	37	0.389	0.537	0.020	0.014
	Score		0.0008	−0.0006	0.0006	0.0004
4.	4gpd	20	0.373	0.509	0.025	0.012
	1a7k	37	0.380	0.510	0.018	0.009
	1gad	37	0.373	0.507	0.011	0.009
	1gd1	55	0.379	0.517	0.013	0.007
	1cer	72	0.371	0.498	0.009	0.009
	1hdg	82	0.379	0.503	0.015	0.011
	Score		−0.0003	0.0014	0.0015	0.0001
5.	1b8p	4	0.384	0.528	0.009	0.012
	5mdh	37	0.383	0.531	0.020	0.010
	1bmd	72	0.387	0.524	0.015	0.009
	Score		−0.0004	0.0006	−0.0009	0.0004
6.	1aw2	15	0.379	0.508	0.010	0.005
	7tim	27	0.387	0.533	0.013	0.004
	1hti	37	0.380	0.521	0.013	0.009
	1tph	37	0.377	0.516	0.014	0.009
	1tcd	37	0.381	0.511	0.014	0.008
	1tre	37	0.386	0.521	0.009	0.004
	1ydv	37	0.365	0.504	0.013	0.006
	1amk	37	0.385	0.524	0.011	0.008
	1tpf	41	0.377	0.511	0.012	0.005
	2btm	55	0.402	0.544	0.009	0.005
	1b9b	82	0.383	0.526	0.014	0.006
	Score		−0.0014	−0.0028	−0.0002	0.0000
7.	2tbs	4	0.390	0.526	0.022	0.015
	1fni	37	0.370	0.508	0.019	0.014
	3tgi	37	0.378	0.523	0.021	0.015
	1trn	37	0.386	0.536	0.015	0.012
	1c1t	37	0.381	0.525	0.021	0.015
	Score		0.0034	0.0009	0.0009	0.0003
	Average score		−0.0005	−0.0003	0.0002	0.0000
	Z-score		−1.4	0.5	1.1	−0.4

[a]Numbers refer to Table I.

among secondary structure elements and/or between molecular domains and enhance the flexibility. The average change rate of H-bonds per residue over all the families is not significant (Z-score = −1.0). Apparently, the decrease of the number of H-bonds is mainly achieved through the alteration of occurrence of H-bond donors and acceptors, rather than by modification of their degree of saturation.

The change of ion pair number per subunit and per residue displays a Z-score equal to −2.5 and −2.2, respectively. These results suggest that variation of ion pairs is a major component of the strategy of the adaptation to extreme temperatures, with progressive decrease toward cold temperatures. The α-amylase family shows the highest decreasing rate in ion pairs, both in the per-subunit and per-residue statistic. Probably, the high reduction of ion pair number in this family may counterbalance the stabilizing effect of the rise in polar accessible surface fraction (vide infra and Table III). The role of ion pairs in

the cold adaptation was also highlighted by single-family comparisons,[1] which showed a reduced Arg/(Arg+Lys) content of the psychrophilic enzymes, and by our previous sequence-based analysis.[11]

Apolar atoms on the surface desolvate the protein in solution. Therefore, reduction of the number of surface-charged residues and increase of apolar surface fraction decrease the number of polar interaction with the solvent and may enhance the flexibility of the protein external shell.[12] The correlation between living temperature and fraction of apolar exposed surface indicates a positive trend though statistically not significant (Z-score = 1.4). Instead, the decrement of fractional accessible surface of side-chain atoms has a Z-score equal to −2.0. This suggests that the increment of apolar exposed surface is mainly achieved by the substitution of exposed side-chains with shorter ones and/or by residue deletions (Fig. 2). In fact, an earlier sequence-based analysis[11] detected the substitutions Glu→Ala, Arg→Lys, Lys→Ser, Lys→Asn, and Arg→Ser from thermo- and mesophiles to psychrophiles at exposed sites.

The higher fraction of the apolar molecular surface and the lower accessible area in cold adapted enzymes may prevent the effects of cold denaturation that involves more favourable interactions of water with polar and apolar atoms at low temperatures.[56,57] In fact, the total accessible surface tends to decrease toward lower temperatures (Table III), although with no statistical significance (Z-score = −1.1).

The temperature corresponding to the maximum stability of a protein increases with increasing hydrophobicity of the interior, whereas the strength of that stability decreases.[58] Indeed, proteins from thermophiles do not display more apolar interiors than proteins from mesophiles.[59,60] Comparative analysis of 10 different triosephosphate isomerase structures from meso- and thermophiles[61] showed that the amount of buried hydrophobic area in isolated subunits is a constant fraction of the total area buried on folding. Even though the average rate for the fraction of apolar buried surface is apparently modest in our data set (Table III), it is significant (Z-score = −2.3). The α-amylase, citrate synthase, and triosephosphate isomerase families display the highest rates. The first two families contain the members with the largest molecular mass per subunit. Larger structural domains have more buried surface area per residue than smaller ones; therefore, large proteins can construct a sufficient extensive and stabilizing hydrophobic core, which gives more "freedom" to the variation of the fraction of apolar surface. Modulation of apolar content may be used to satisfy different requirements of structural flexibility at different living temperatures.

Apolar interactions in the core of a protein are generally considered as a guiding force in determining protein folding and stability,[53,54] and generally they increase with increasing host temperature.[58] Therefore, it would be expected that the loss of such interactions could result in a lower rigidity of the molecular core in cold adapted enzymes. Although the variation of the fraction of apolar contact surface with environmental temperature is negative in six of seven families, it is not statistically significant (Z-score = −1.5). Consequently, the real contribution of this factor to the adaptation to low temperatures is unclear.

Nonetheless, the overview of surface properties (fraction apolar exposed, apolar buried, and apolar contact surfaces) suggests that cold adaptation could entail a subtle reorganization of the total exposed and buried surfaces, with a general reduction of hydrophobic contact area. It is of interest that our previous sequence-based analysis[11] detected the exchanges Val→Ala and Val→Ile, the former more significant, occurring in the hydrophobic core. In fact, the combined effect of these substitutions is the reduction of the number of core carbon atoms from six to five, with overall loss of atomic surface available to hydrophobic contacts.

To assess the impact of the variation of surface properties on protein stability, the change in solvation free energy from unfolded to folded conformation (ΔG_s) at room temperature was predicted. The higher the ΔG_s, the lower the stability in solution. Therefore, positive differences between the ΔG_s of the cold and the warm enzymes indicate that the former is less stable than the latter in solution at room temperature. Although five of the seven families show an increasing ΔG_s per residue toward low temperatures (Table III), the final Z-score is not significant (1.3). However, an interesting correlation is observed between the family rates of the variation in solvation ΔG_s and the respective variations in the fractional apolar surface, accessible and/or buried. For example, the high positive rate in the GAPDH family corresponds to the highest rate of increase in fractional apolar accessible surface (Table III). The large rates shown by the triosephosphate isomerase and citrate synthase families can be explained by the corresponding reduction in fractional apolar buried surface (Table III). On the contrary, α-amylase family shows a modest increase in solvation free energy (Table III). This can result from the balance between the effects of the great negative rates of variation in fractional apolar buried and exposed surfaces. However, caution should be played in the interpretation of such results because of the incompleteness of current knowledge about the various energetic contributions to protein folding as well as the controversy surrounding solvation energy functions on which the ΔG_s calculation performed here depends.

The analysis of the change in secondary structural content does not highlight any significant variation. Elastase and trypsin families contain only a little fraction of helices, so the net score found for this conformation could be biased by the family differences in secondary structural composition. It is of interest that the same analysis suggests that deletions occur more frequently with decreasing temperatures (Z-score = 2.0), mostly at nonrepetitive regions. This can contribute to the observed trend of reduction in total solvent-exposed surface of cold adapted enzymes (Table III).

Globular proteins are generally characterized by optimal packing of their interior,[62] which does not exclude the presence of imperfections and cavities.[63] Comparative

analyses of protein structures revealed a correlation between reduction of cavity volume and thermal stability of thermophilic proteins.[64,65] Indeed, despite very small cavity volumes,[63] their contribution to the molecular energy is significant and can influence protein stability.[66,67] Less clear is the relationship between packing and stability at high temperatures. Only in a few cases some correlation was detected.[51,68–70] In our sample, the average rate of variation of packing density is not significant (Z-score = −1.4), suggesting that this feature does not represent a general strategy of adaptation to low temperatures. Packing density of buried residues is even less variable (Table IV). It should be emphasized that the analysis of average packing density cannot detect local perturbation of packing, which may be of functional relevance. As for packing, the analysis of variation in fractional volume of cavities does not show any significant correlation with the adaptation to low temperatures (Z-score = 1.1). Water molecules contained in the cavities[71,72] were also considered. Although the impact of water molecules in the interior of a protein on the stability is still debated, experimental studies showed that their presence in cavities created by site-directed mutagenesis contributes favorably to the stability[73,74] and that their free energy defines them as integral parts of the structure.[43] The calculation of water content in the enzymes of our data set does not reveal significant variations (Z-score = −0.4) with decreasing temperatures (Table IV), indicating that buried water does not seem correlated to cold adaptation.

It may be argued that the optimum growth temperatures listed in Table I are only indicative. Except for homeotherms and for species living within homeothermic hosts, the other organisms (salmon, lobster, marine bacteria, and meal worm) may be exposed to a range of temperatures. To assess the impact of variable living temperatures on the calculated rates of structural features, a statistical test was performed. Temperatures for all the organisms with potentially variable living temperatures were randomly modified within a range of ±5°C around the indicative temperature, and the corresponding Z-score was recalculated. Assignment of randomly changed temperatures and determination of the Z-score as described in Material and Methods were repeated 1000 times. All the significant scores were confirmed. Z-scores for the change in ion pairs per subunit and per residue range from −2.4 to −3.4 (mean = −2.9) and from −2.0 to −3.1 (mean = −2.6), respectively. The contribution of side-chain atoms to the exposed surface has a Z-score varying from −1.8 to −2.9 (mean = −2.3). The Z-score for the variation in apolar fraction of the buried surface ranges from −2.1 to −3.1 (mean = −2.6). H-bond rate per subunit displays a Z-score ranging from −1.7 to −2.6 (mean = −2.1). Consequently, it cannot be definitively excluded that variation of H-bonds plays a role in cold adaptation.

In conclusion, we analyzed the variations among seven protein families of several structural properties thought to be related to stability; of these properties, only the number of ion pair interactions, the fraction of accessible surface of side-chains, and the fraction of apolar occluded surface appeared to be significantly correlated to the adaptation to low temperatures, because all of them tend to decrease with the organism optimal temperature. Other properties (e.g., fractional apolar contact surface and packing density) display trends that may be relevant to adaptation to low temperatures for single families, although statistically not significant on average for the entire data set. Our results, if considered on an intrafamily basis, can give insights into family-specific strategies of adaptation. The conclusions reached by our study can help the interpretation of the peculiarities of the enzyme activity at low temperatures and can indicate strategies for rational enzyme engineering.

ACKNOWLEDGMENTS

This work will be submitted by G. G. in partial fulfillment of the requirements of the degree of Dottorato di Ricerca at the Università di Roma La Sapienza. The authors thank the two anonymous referees for their helpful and constructive criticism.

REFERENCES

1. Feller G, Gerday C. Psychrophilic enzymes: molecular basis of cold adaptation. Cell Mol Life Sci 1997;53:830–841.
2. Gerday C, Aittaleb M, Arpigny JL, Baise E, Chessa JP, Garsoux G, Petrescu I, Feller G. Psychrophilic enzymes: a thermodynamic challenge. Biochim Biophys Acta 1997;1342:119–131.
3. Gerday C, Aittaleb M, Bentahir M, Chessa JP, Claverie P, Collins T, D'Amico S, Dumont J, Garsoux G, Georlette D, Hoyoux A, Lonhienne T, Meuwis MA, Feller G. Cold-adapted enzymes: from fundamentals to biotechnology. Trends Biotech 2000;18:103–107.
4. Marshall CJ. Cold-adapted enzymes. Trends Biotech 1997;15:359–364.
5. Russell NJ. Molecular adaptations in psychrophilic bacteria: potential for biotechnological applications. Adv Biochem Eng Biotech 1998;61:1–21.
6. Vogt G, Woell S, Argos P. Protein thermal stability, hydrogen bonds, and ion pairs. J Mol Biol 1997;269:631–643.
7. Szilágyi A, Závodszky P. Structural differences between mesophilic, moderately thermophilic and extremely thermophilic protein subunits: results of a comprehensive survey. Structure 2000; 8:493–504.
8. Wallon G, Lovett ST, Magyar C, Svingor A, Szilágyi A, Závodszky P, Ringe D, Petsko GA. Sequence and homology model of 3-isopropylmalate dehydrogenase from the psychrotrophic bacterium *Vibrio* sp. I5 suggest reasons for thermal instability. Protein Eng 1997;10: 665–672.
9. Narinx E, Baise E, Gerday C. Subtilisin from psychrophilic antarctic bacteria: characterization and site-directed mutagenesis of residues possibly involved in the adaptation to cold. Protein Eng 1997;10:1271–1279.
10. Aittaleb M, Hubner R, Lamotte-Brasseur J, Gerday C. Cold adaptation parameters derived from cDNA sequencing and molecular modelling of elastase from Antarctic fish *Notothenia neglecta.* Protein Eng 1997;10:475–477.
11. Gianese G, Argos P, Pascarella S. Structural adaptation of enzymes to low temperatures. Protein Eng 2001;14:141–148.
12. Aghajari N, Feller G, Gerday C, Haser R. Structures of the psychrophilic *Alteromonas haloplanctis* α-amylase give insights into cold adaptation at a molecular level. Structure 1998;6:1503–1516.
13. Russell RJM, Gerike U, Danson MJ, Hough DW, Taylor GL. Structural adaptation of the cold-active citrate synthase from an Antarctic bacterium. Structure 1998;6:351–361.
14. Berglund GI, Smalås AO, Outzen H, Willassen NP. Purification and characterization of pancreatic elastase from North Atlantic salmon (*Salmo salar*). Mol Mar Biol Biotech 1998;7:105–114.
15. Smalås AO, Heimstad ES, Hordvik A, Willassen NP, Male R. Cold adaption of enzymes: structural comparison between salmon and bovine trypsins. Proteins 1994;20:149–166.
16. Murthy MR, Garavito RM, Johnson JE, Rossmann MG. Structure

of lobster apo-D-glyceraldehyde-3-phosphate dehydrogenase at 3.0 Å resolution. J Mol Biol 1980;138:859–872.
17. Tanner JJ, Hecht RM, Krause KL. Determinants of enzyme thermostability observed in the molecular structure of *Thermus aquaticus* D-glyceraldehyde-3-phosphate dehydrogenase at 2.5 Å resolution. Biochemistry 1996;35:2597–2609.
18. Kim S-Y, Hwang KY, Kim S-H, Sung H-C, Han YS, Cho Y. Structural basis for cold adaptation. J Biol Chem 1999;271:11761–11767.
19. Alvarez M, Zeelen JP, Mainfroid V, Rentier-Delrue F, Martial JA, Wyns L, Wierenga RK, Maes D. Triose-phosphate isomerase (TIM) of the psychrophilic bacterium *Vibrio marinus.* J Biol Chem 1998;273:2199–2206.
20. Villeret V, Chessa J-P, Gerday C, Van Beeumen J. Preliminary crystal structure determination of the alkaline protease from the Antarctic psychrophile Pseudomonas aeruginosa. Protein Sci 1997; 6:2462–2464.
21. Sussman JL, Lin D, Jiang J, Manning NO, Prilusky J, Ritter O, Abola EE. Protein Data Bank (PDB): database of three-dimensional structural information of biological macromolecules. Acta Crystallogr D Biol Crystallogr 1998;54:1078–1084.
22. Pearson WR, Lipman DJ. Improved tools for biological sequence comparison. Proc Natl Acad Sci USA 1988;85:2444–2448.
23. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 1993;26:283–291.
24. Hooft RW, Vriend G, Sander C, Abola EE. Errors in protein structures. Nature 1996;381:272.
25. Shindyalov IN, Bourne PE. Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. Protein Eng 1998;11:739–747.
26. Insight II, Version 98.0, Molecular Modeling System, User Guide. San Diego: Biosym/MSI; 1998.
27. Querol E, Perez-Pons JA, Mozo-Villarias A. Analysis of protein conformational characteristics related to thermostability. Protein Eng 1996;9:265–271.
28. McDonald IK, Thornton JM. Satisfying hydrogen bonding potential in proteins. J Mol Biol 1994;238:777–793.
29. Baker EN, Hubbard RE. Hydrogen bonding in globular proteins. Prog Biophys Mol Biol 1984;44:97–179.
30. Barlow DJ, Thornton JM. Ion-pairs in proteins. J Mol Biol 1983;168:867–885.
31. Musafia B, Buchner V, Arad D. Complex salt bridges in proteins: statistical analysis of structure and function. J Mol Biol 1995;254: 761–770.
32. Naccess, version 2.1.1, Computer Program, Hubbard SJ, Thornton JM. London: Department of Biochemistry and Molecular Biology, University College; 1993.
33. Pattabiraman N, Ward KB, Fleming PJ. Occluded molecular surface: analysis of protein packing. J Mol Recognit 1995;8:334–344.
34. Connolly ML. Solvent-accessible surfaces of proteins and nucleic acids. Science 1983;221:709–713.
35. Juffer AH, Eisenhaber F, Hubbard SJ, Walther D, Argos P. Comparison of atomic solvation parametric sets: applicability and limitations in protein folding and binding. Protein Sci 1995;4:2499–2509.
36. Eisenberg D, Wesson M, Yamashita M. Interpretation of protein folding and binding with atomic solvation parameters. Chem Scrip 1989;29A:217–221.
37. Eisenhaber F, Argos P. Improved strategy in analytic surface calculation for molecular systems: handling of singularities and computational efficiency. J Comp Chem 1993;14:1272–1280.
38. Eisenhaber F, Lijnzaad P, Argos P, Sander C, Scharf M. The double cubic lattice method: efficient approaches to numerical integration of surface area and volume and to dot surface contouring of molecular assemblies. J Comp Chem 1995;16:273–284.
39. Frishman D, Argos P. Knowledge-based protein secondary structure assignment. Proteins 1995;23:566–579.
40. Pascarella S, De Persio R, Bossa F, Argos P. Easy method to predict solvent accessibility from multiple protein sequence alignments. Proteins 1998;32:190–199.
41. Nicholls A, Sharp KA, Honig B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins 1991;11:281–296.
42. Liang J, Edelsbrunner H, Woodward C. Anatomy of protein pockets and cavities: measurement of binding site geometry and implications for ligand design. Protein Sci 1998;7:1884–1897.
43. Zhang L, Hermans J. Hydrophilicity of cavities in proteins. Proteins 1996;24:433–438.
44. Hubbard SJ, Gross K-H, Argos P. Intramolecular cavities in globular proteins. Protein Eng 1994;7:613–626.
45. Fleming PJ, Richards FM. Protein packing: dependence on protein size, secondary structure and amino acid composition. J Mol Biol 2000;299:487–498.
46. Serrano L, Fersht AR. Capping and alpha-helix stability. Nature 1989;342:296–299.
47. Sheridan PP, Panasik N, Coombs JM, Brenchley JE. Approaches for deciphering the structural basis of low temperature enzyme activity. Biochim Biophys Acta 2000;1543:417–433.
48. Russell RB, Barton GJ. Structural features can be unconserved in proteins with similar folds. An analysis of side-chain to side-chain contacts secondary structure and accessibility. J Mol Biol 1994;244: 332–350.
49. Goldman A. How to make my blood boil. Structure 1995;3:1277–1279.
50. Myers JK, Pace CN. Hydrogen bonding stabilizes globular proteins. Biophys J 1996;71:2033–2039.
51. Vogt G, Argos P. Protein thermal stability: hydrogen bonds or internal packing? Fold Des 1997;2:S40–S46.
52. Xiao L, Honig B. Electrostatic contributions to the stability of hyperthermophilic proteins. J Mol Biol 1999;289:1435–1444.
53. Dill KA. Dominant forces in protein folding. Biochemistry 1990;29: 7133–7155.
54. Pace CN, Shirley BA, McNutt M, Gajiwala K. Forces contributing to the conformational stability of proteins. FASEB J 1996;10:75–83.
55. Fontana A, De Filippis V, de Laureto PP, Scaramella E Zambonin M. Rigidity of thermophilic enzymes. In: Ballesteros A, Plou FJ, Iborra JL, Halling PJ, editors. Stability and stabilization of biocatalysts. Amsterdam: Elsevier Science; 1998. p 277–294.
56. Makhatadze GI, Privalov PL. Energetics of protein structure. Adv Protein Chem 1995;47:307–425.
57. Graziano G, Catanzano F, Riccio A, Barone G. A reassessment of the molecular origin of cold denaturation. J Biochem 1997;122:395–401.
58. Creighton TE. Stability of folded conformations. Curr Opin Struct Biol 1991;1:5–16.
59. Frömmel C, Sander C. Thermitase, a thermostable subtilisin: comparison of predicted and experimental structures and the molecular cause of thermostability. Proteins 1989;5:22–37.
60. Backmann J, Schafer G, Wyns L, Bonisch H. Thermodynamics and kinetics of unfolding of the thermostable trimeric adenylate kinase from the archaeon *Sulfolobus acidocaldarius.* J Mol Biol 1998;284:817–833.
61. Maes D, Zeelen JP, Thanki N, Beaucamp N, Alvarez M, Thi MHD, Backmann J, Martial JA, Wyns L, Jaenicke R, Wierenga RK. The crystal structure of triosephosphate isomerase (TIM) from *Thermotoga maritima:* a comparative thermostability structural analysis of ten different TIM structures. Proteins 1999;37:441–453.
62. Richards FM. Areas, volumes, packing and protein structure. Annu Rev Biophys Bioeng 1977;6:151–176.
63. Rashin AA, Iofin M, Honig B. Internal cavities and buried waters in globular proteins. Biochemistry 1986;25:3619–3625.
64. Russell RJM, Hough DW, Danson MJ, Taylor GL. The crystal structure of citrate synthase from the thermophilic archaeon, *Thermoplasma acidophilum.* Structure 1994;2:1157–1167.
65. Knapp S, de Vos WM, Rice D, Ladenstein R Crystal structure of glutamate dehydrogenase from the hyperthermophilic eubacterium *Thermotoga maritima* at 3 Å resolution. J Mol Biol 1997;267: 916–932.
66. Eriksson AE, Baase WA, Zhang XJ, Heinz DW, Blaber M, Baldwin EP, Matthews BW. Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. Science 1992;255:178–183.
67. Ishikawa K, Okumura M, Katayanagi K, Kimura S, Kanaya S, Nakamura H, Morikawa K. Crystal structure of ribonuclease H from *Thermus thermophilus* HB8 refined at 2.8 Å resolution. J Mol Biol 1993;230:529–542.
68. Britton KL, Baker PJ, Borges KM, Engel PC, Pasquo A, Rice DW, Robb FT, Scandurra R, Stillman TJ, Yip KS. Insights into thermal stability from a comparison of the glutamate dehydrogenase from *Pyrococcus furiosus* and *Thermococcus litoralis.* Eur J Biochem 1995;229:688–695.
69. Yip KS, Stillman TJ, Britton KL, Artymiuk PJ, Baker PJ, Sedelnikova SE, Engel PC, Pasquo A, Chiaraluce R, Consalvi V.

The structure of *Pyrococcus furiosus* glutamate dehydrogenase reveals a key role for ion-pair networks in maintaining enzyme stability at extreme temperatures. Structure 1995;3:1147–1158.
70. Karshikoff A, Ladenstein R. Proteins from thermophilic and mesophilic organisms essentially do not differ in packing. Protein Eng 1998;11:867–872.
71. Kossiakoff AA, Sintchak MD, Shpungin J, Presta LG. Analysis of solvent structure in proteins using neutron D2O-H2O solvent maps: pattern of primary and secondary hydration of trypsin. Proteins 1992;12:223–236.
72. Ernst JA, Clubb RT, Zhou H-X, Gronenborn AM, Clore M. Demonstration of positionally disordered water within a protein hydrophobic cavity by NMR. Science 1995;267:1813–1817.
73. Takano K, Funahashi J, Yamagata Y, Fujii S, Yutani K. Contribution of water molecules in the interior of a protein to the conformational stability. J Mol Biol 1997;274:132–142.
74. Funahashi J, Takano K, Yamagata Y, Yutani K. Contribution of amino acid substitutions at two different interior positions to the conformational stability of human lysozyme. Protein Eng 1999;12: 841–850.