Journal of Theoretical Biology 383 (2015) 130–137

Contents lists available at ScienceDirect

Journal of Theoretical Biology

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

Protein cold adaptation: Role of physico-chemical parameters in adaptation of proteins to low temperatures

Soheila Shokrollahzade [a], Fatemeh Sharifi [b], Akbar Vaseghi [c], Maryam Faridounnia [d], Samad Jahandideh [e,*]

[a] Department of Medicinal Biotechnology, Iran University of Medical Sciences, Tehran, Iran
[b] Department of Computer Science, Shahid Beheshti University, G.C., Tehran, Iran
[c] Young Researchers and Elite club, Ardabil Branch, Islamic Azad University, Ardabil, Iran
[d] Bijvoet Center for Biomolecular Research, Utrecht University, Utrecht, The Netherlands
[e] Program on Bioinformatics and Systems Biology, Sanfore Burnham Prebys Medical Discovery Institute, 10901 N Torrey Pines Rd, La Jolla, CA 92307, United States

HIGHLIGHTS

- Psychrophilicity rules are not merely the inverse rules of thermostability.
- Psychrophilic/mesophilic protein pairs are different in terms of physico-chemical properties.
- Contact orders of psychrophilic/mesophilic protein pairs are not different.

ARTICLE INFO

Article history:
Received 18 March 2015
Received in revised form
21 June 2015
Accepted 16 July 2015
Available online 29 July 2015

Keywords:
Structural features
Thermostability
Systematics analysis
Packing density
Surface accessibility

ABSTRACT

During years 2007 and 2008, we published three papers (Jahandideh, 2007a, JTB, 246, 159-166; Jahandideh, 2007b, JTB, 248, 721-726; Jahandideh, 2008, JTB, 255, 113-118) investigating sequence and structural parameters in adaptation of proteins to low temperatures. Our studies revealed important features in cold-adaptation of proteins. Here, we calculate values of a new set of physico-chemical parameters and perform a comparative systematic analysis on a more comprehensive database of psychrophilic–mesophilic homologous protein pairs. Our obtained results confirm that psychrophilicity rules are not merely the inverse rules of thermostability; for instance, although contact order is reported as a key feature in thermostability, our results have shown no significant difference between contact orders of psychrophilic proteins compared to mesophilic proteins. We are optimistic that these findings would help future efforts to propose a strategy for designing cold-adapted proteins.

1. Introduction

Living organisms, based on the temperature they endure, are categorized into psychrophiles, mesophiles and thermophiles. Psychrophiles and thermophiles are those that are adapted to regions with extreme temperature, where only few organisms could survive (Baldasseroni and Pascarella, 2009). By using physiological strategies, psychrophiles grow in 5 °C in the depth of oceans, in glaciers and mountains, and maintain their own function in these temperatures (Gianese et al., 2002). Near three-quarters of the Earth's surface is owned by cold regions (Feller, 2013). Psychrophilic proteins are found in prokaryotic and eukaryotic organisms including bacteria, archaea, algae, and yeasts as well as glaciers ice worms, plants and animals (Buzzini et al., 2012; Cameron et al., 2012; Cavicchioli, 2006). From the first time in 1975 that Morita introduced psychrophilic bacteria, molecular bases of cold adaptation have attracted the interest of researchers (Cipolla et al., 2012; Morita, 1975).

From the industrial and biotechnological point of view, psychrophilic proteins are endowed with many considerable merits, which make them applicable. Psychrophilic proteins save energy due to decrease in the free energy of activation $\Delta G^{\#}$. Consequently, less energy is required to inactivate a psychrophilic enzyme and therefore it can increase the reaction speed (Baldasseroni and Pascarella, 2009; D'Amico et al., 2002). This makes psychrophilic proteins economically beneficial for the processes in which they are involved. Examples of such industries include detergent, food, textile manufacturing, pharmaceutical, biofuels, and energy production (Feller, 2013; Jahandideh et al., 2007a).

* Corresponding author.
E-mail address: sjahandideh@burnham.org (S. Jahandideh).

http://dx.doi.org/10.1016/j.jtbi.2015.07.013
0022-5193/

Structural and sequential properties of cold-adapted proteins are subjected to many studies either as small datasets of individual proteins or large ones such as genomics, metagenomics and proteomics studies (Goodchild et al., 2004; Grzymski et al., 2006; Lauro et al., 2009; Metpally and Reddy, 2009). To address some of those important studies, comparative genomics studies for amino acid composition in some cold adapted organisms have shown higher content of noncharged polar amino acids and a lower content of hydrophobic amino acids as well as more propensity in solvent accessible area for some hydrophobic residues (Casanueva et al., 2010; Saunders et al., 2003). In addition, many researches have reported the difference in occurrence of some amino acid substitutions in psychrophilic proteins as a likely adaptation strategy by psychrophiles (Jahandideh et al., 2008; Methe et al., 2005). As for structural characterization of psychrophilic proteins and their pertinent adaptation process, cavities and clefts are indispensable factors to consider in cold-adaptation (Paredes et al., 2011). Studies in cavities, clefts and void volumes by many methods have suggested larger cavities in psychrophilic proteins (Collins et al., 2002; Feller, 2013; Liang et al., 1998; Paredes et al., 2011).

Up to now, many efforts have been done by previous investigators in using physico-chemical parameters and properties although for different biomacromolecules and purposes (Chen et al., 2013, 2014; Feng et al., 2013; Guo et al., 2014; Kandaswamy et al., 2011; Lin et al., 2014; Liu et al., 2015a, 2015b; Qiu and Xiao, 2014; Xiao and Wang, 2012, 2013; Xiao and Lin, 2013; Xiao et al., 2013a, 2013b; Wu and Xiao, 2010). In this work, we have investigated physico-chemical properties of mesophilic and psychrophilic protein pairs. For this purpose, we have created a dataset of 30 pairs of mesophilic and psychrophilic proteins. To investigate the differences, we compared several structural properties; molecular weight (MW) per residues, radius of gyration (ROG), hydrogen bond (HB), packing density (PD), accessible surface area (ASA), buried surface area (BSA), and number of cavities. These global parameters are assigned to general structures of psychrophilic/mesophilic pairs, while in our previous studies, local parameters attributed to specific locations of proteins have been discussed (Jahandideh et al., 2007a, 2007b, 2008). The results suggest some possible general rules for protein design experiments aimed to produce enzymes catalytically more effective at low temperatures.

2. Materials and methods

2.1. Database

We followed some steps to provide a psychrophilic–meshophilic homologus database of proteins. Initially, keywords concerning cold adaptation were used to search for psychrophilic proteins in protein data bank (PDB) (Berman et al., 2000). To ensure that psychrophilic proteins are correctly chosen, related research articles reporting the structures were reviewed. Then, non-wild-type proteins, i.e. mutant proteins, and those composed of abnormal residues, such as Asx or Glx, were eliminated from the database. This provided us with a list of 30 mesophilic proteins. Finally, protein–protein BLAST (blastp) was used, by entering each single psychrophilic protein's PDB ID as a query, to find its best homologous mesophilic protein pair with more than 25% identity. Therefore, in the blast result, subjects with the most identity were selected and their accession number was checked in protein data bank. To make sure the selected protein is mesophilic, molecular description including its classification, organism as well as habitat was checked. All of the selected homologous proteins were in the same protein and enzyme classification as their psychrophilic homologs. The database contains 30 pairs of psychrophilic–meshophilic homologous proteins, all of which are X-ray structures determined with a resolution of at least 3 Å. PDB IDs, species and resolution of these protein structures are shown in Table 1.

2.2. Data analysis

2.2.1. Packing density

One of the structural features denoting the intrinsic compressibility of proteins is *packing density* (PD). Packing density provides information associated with stability and flexibility of proteins. A distinct example of this correlation is hydrophobic interactions, which, by bunching hydrophobic residues tightly together, make the protein more thermostable (Criswell et al., 2003).

Structural analysis has shown that adaptation of hyper-thermophiles is strongly related to increase in *packing density* in the surface of protein due to decrease in the number of cavities (Baldasseroni and Pascarella, 2009). Pack et al. have shown different packing density in exposed surface of thermophilic proteins (Pack and Yoo, 2005). To address the question of whether or not the differences in cold adaptation of psychrophilic proteins is due to the *packing density*, we calculated *packing density* for each particular pair of PDB structures by Voronoia software. Voronoia calculate radius of atoms for amino acids of intended proteins. The atomic packing density is defined as

$$\mathrm{PD} = \frac{V(\mathrm{VdW})}{[V(\mathrm{VdW}) + V(\mathrm{SE})]}$$

where V(Vdw) is volume of van der Waals radius and V(SE) is surface excluded volume voronoia attribute van der Waals radii for neighboring atoms in the proximity of a specific residue. It then splits the solvent excluded protein by a polyhedron space, called voronoia volume (Goede et al., 1997).

2.2.2. Cavity

A cavity is often a buried region in protein without opening to outside (Fig. 1). It acts as ligand-binding site or any putative internal water and plays important roles in binding and flexibility of proteins (Liang et al., 1998). Some studies have reported larger cavities in psychrophilic proteins. Larger cavities, by adopting more water molecules within themselves, make it possible to act in low temperature (Paredes et al., 2011). Solvent probes with often a specific size are used to measure cavity volume and once the cavity is filled the probe cannot measure the cavity. In addition to volume of cavities, numbers of cavities in proteins are important. We used the web interface of Voronoia program to examine cavity volumes for each pair of proteins. This program (available at http://bioinformatics.charite.de/voronoia/) uses a 1.4 Å solvent probe to examine cavities (Goede et al., 1997).

2.2.3. Surface area

The total surface area of a biomolecule that is approachable by a solvent is called accessible surface area (ASA). In 1971, it was introduced as the area of the region bounded by tracing the locus of solvent probe, as it is rolling over the van der Waals surface (Fig. 1) by Lee and Richard introduced this concept as the area of the region bounded by tracing the locus of solvent probe, as it is rolling over the van der Waals surface (Fig. 1) (Lee and Richards, 1971). This boundary is somehow equivalent to extending the surfaces of atoms with a radius equivalent to solvent probe radius. These expanded spheres might intersect; the surface exterior to the union of these overlapping spheres is considered (Richmond, 1984). Shrake–Rupley algorithm, a common and simple algorithm, uses a numerical method, and a solvent probe radius of 1.4 Å to calculate ASA. Approximately equal to the radius of water molecule, 1.4 Å is a typical value for probe radius, however, different algorithms may consider different probe radii (Shrake and Rupley, 1973).

ASA values impact many characteristics of a protein, e.g. thermostability and flexibility. Therefore, they have been used in different articles to facilitate the process of solving hydrophobicity effect problem, and to estimate free energy, heat capacity of hydration, and enthalpy (Marsh, 2013; Ooi et al., 1987; Richmond, 1984; Stellwagen and Wilgus, 1978). By using PDBePISA online server

(Krissinel and Henrick, 2007), we extracted ASA values of two groups of mesophilic and psychrophilic proteins. To increase the reliability of the dataset, these ASA values were cross-checked with those from DSSP database. Since generally proteins with heavy molecular weight have high ASA values, these ASA values were divided by their corresponding molecular weight to obtain unbiased data. Since generally proteins with heavy molecular weight have high ASA values, these ASA values were divided by their corresponding molecular weight to obtain unbiased data (Janin et al., 1988). In addition, using the aforementioned approach, we extracted buried surface area at the interface of protein–protein complexes.

2.2.4. Buried surface area

Buried surface area (BSA) measures the size of the interface in a protein–protein complex, that is, the sum of subunit accessible surface areas that are not accessible anymore when a complex is formed. Therefore, for the monomers, BSA is equal to zero. BSA has

Table 1
The psychrophilic–mesophilic homologous database.

Protein Name	Source	Species	PDB code	Resolution (Å)
Alpha-amylase	Psychrophilic	*Alteromonas haloplanctis*	1AQH	2
	Mesophilic	Barley seed	1AVA/C	1.9
Cellulase Cel5	Psychrophilic	*Pseudoalteromonas haloplanktis*	1TVN/A	1.41
	Mesophilic	*Erwinia cherysanthemi*	1EGZ/C	2.3
Iron superoxide dismutase	Psychrophilic	*Pseudoalteromonas haloplanktis*	3LJF/A	2.1
	Mesophilic	*T. thermophilus*	1ISA/A	1.8
S-formylglutathione hydrolase	Psychrophilic	*Pseudoalteromonas haloplanktis* TAC125	3LS2/A	2.1
	Mesophilic	*Agrobacterium tumefaciens*	3E4D/A	2.01
Alkaline protease	Psychrophilic	Pseudomonas tac Ii 18	1G9K	1.96
	Mesophilic	*Pseudomonas aeruginosa*	1KAP	1.64
Beta-lactamase	Psychrophilic	*Pseudomonas fluorescens*	2QZ6	2.26
	Mesophilic	*Pseudomonas aeruginosa*	2WZX	1.4
Citrate synthase	Psychrophilic	Antarctic bacterium DS2-3R	1A59	2.09
	Mesophilic	*Sulfolbus tokodaii* strain7	1VGP/A	2.7
Beta-galactosidase	Psychrophilic	Arthrobacter sp. C2-2 (isoenzyme C2-2-1)	1YQ2/A	1.9
	Mesophilic	*Escherchia coli*	1BGL/H	2.5
Alkaline phosphatase	Psychrophilic	Antarctic bacterium Tab5	2IUC/A	1.95
	Mesophilic	*Escherichia coli*	1Y7A/A	1.77
Aminopeptidase	Psychrophilic	*Colwellia psychrerythraea*	3CIA/A	2.7
	Mesophilic	*Thermoplasma acidophilum*	1Z1W/A	2.7
Phenylalanine hydroxylase	Psychrophilic	*Colwellia psychrerythraea* 34h	2V27	1.5
	Mesophilic	*Chromobacterium violaceum*	3TK2	1.35
Malate dehydrogenase	Psychrophilic	*Aquaspirillum arcticum*	1B8P	1.9
	Mesophilic	*Burkholderia pseudomallei*	3D5T/A	2.51
Aspartate transcarbamoylase	Psychrophilic	*Moritella profunda*	2BE7/A	2.85
	Mesophilic	*Escherichia coli*	1EKX/A	1.95
Adenylate kinase	Psychrophilic	*Marinibacillus marinus*	3FB4	2
	Mesophilic	*Bacillus globisporus*	1S3G	2.25
Triosephosphate isomerase	Psychrophilic	*Vibrio marinus*	1AW2/A	2.65
	Mesophilic	*Escherichia coli*	4IOT/A	1.85
Catalase	Psychrophilic	*Vibrio salmonicida*	2ISA/A	1.97
	Mesophilic	*Pseudomonas aeruginosa*	4E37/A	2.53
Endonuclease	Psychrophilic	*Vibrio salmonicida*	2PU3	1.5
	Mesophilic	*Vibrio vulnificus*	1OUP/A	2.3
Lipase	Psychrophilic	*Candida antarctica*	1TCA	1.55
	Mesophilic	*Homo sapiens*	1HLG/B	3
Apo LDH	Psychrophilic	*Champsocephalus gunnari*	2V65/A	2.35
	Mesophilic	*Cyprinus carpio*	1V6A/A	2.3
Trypsin	Psychrophilic	Salmo salar	2TBS	1.8
	Mesophilic	Salmo salar	1BIT	1.83
Elastase	Psychrophilic	Salmo salar	1ELT	1.61
	Mesophilic	Suc scrofa	1EAI	2.4
Glycosylase	Psychrophilic	Atlantic cod	1OKB/B	1.9
	Mesophilic	*Homo sapiens*	3TKB	1.5
Protease	Psychrophilic	Pseudomonas	1G9K	1.96
	Mesophilic	*Serratia marcescens*	1SAT	1.75
Chitinase B	Psychrophilic	Arthrobacter TAD20	1KFW	1.74
	Mesophilic	*Serratia marcescens*	1E15/B	1.9
Hemoglobin	Psychrophilic	Antarctic Fish Pagothenia Bernacchii	1PBX/A	2.5
	Mesophilic	*Leiostomus xanthurus*	1SPG/A	1.95
Beta-glucosidase	Psychrophilic	*Micrococcus antarcticus*	3W53	2.2
	Mesophilic	Streptomyces Sp	1GNX/A	1.68
Leucine dehydrogenase	Psychrophilic	*Sporosarcina psychrophila*	3VPX/A	2.55
	Mesophilic	*Bacillus sphaericus*	1LEH/A	2.2
Tryptophan synthase	Psychrophilic	*Shewanella frigidimarina* K14-2	3VND/A	2.6
	Mesophilic	*Vibrio cholerae* O1 biovar El	3NAV/A	2.1
Glutathione S-transferase	Psychrophilic	*Clamlaternula elliptica*	3QAV	2.1
	Mesophilic	*Arabidopsis thaliana*	1E6B	1.65
Type III antifreeze protein RD1	Psychrophilic	*Lycodichthys dearborni*	1UCS	0.62
	Mesophilic	*Macrozoarces americanus*	1HG7	1.15

an impact on protein (Marsh, 2013). Using the same approach as we described for ASA, BSA data were extracted from PDBePISA server (Krissinel and Henrick, 2007).

2.2.5. Molecular weight per residue

Molecular weight is a critical feature of a protein in terms of biochemical characterization. MW were calculated per residue for all psychrophilic and mesophilic homologous pairs of protein structures. Relative MW per residue data was obtained by using the Compute PI/MW web tool (http://web.expasy.org/compute_pi/).

2.2.6. Radius of gyration (ROG)

Radius of gyration is one of the structural parameters, which is defined as the root-mean-square distance between each atom of the protein and their center of gravity. ROG indicates the compactness, and degree of folding of a protein or polypeptide chain. It means that a protein with a high radius of gyration is packed tightly. It is also informative about native and denatured state of the protein. Studies have shown different radius of gyration in different secondary structures. α proteins have the highest radius of gyration therefore are less packed, while β proteins, (α+β)-proteins and α/β proteins possess lower radius of gyration and tighter packing, respectively (Lobanov et al., 2008). ROG of a protein changes as solvent environment and PH change (Srinivas et al., 2014). ROG can be calculated by experimental techniques such as small-angle X-ray scattering accurately, however, ROG is predictable by using computational methods (Simon, 1971). We used online SCFBio database (http://www.scfbio-iitd.res.in/software/proteomics/rg.jsp) to separately measure gyration radius of each chosen PDB structure.

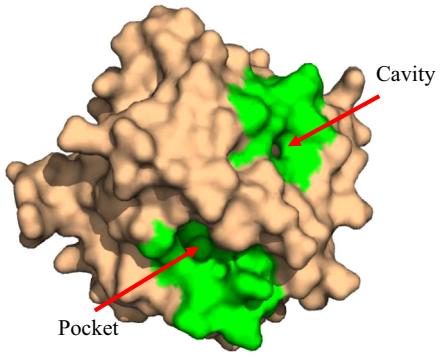

Fig. 1. The structure of adenylate kinases from the psychrophile *Bacillus globisporus* (PDB ID: 1S3G). Arrows show interior cavity and surface pocket it may contain. The structure is shown as surface representation by Pymol software.

2.2.7. Hydrogen bonds

Hydrogen bonds, as one of the most important classes of molecular interactions in biology that confer directionality and specificity to the intermolecular interactions, are of great importance for protein structure and stability (McDonald and Thornton, 1994). In order to identify hydrogen bonds for all psychrophilic and mesophilic proteins in our dataset, we used a program called DSSP. DSSP uses an algorithm that defines secondary structure, geometrical features, and solvent exposure of proteins, based on atomic coordinates of PDB files. Properties of each protein are presented in a Table as a Supplementary file. In order to test normal distribution of properties we applied Lilliefors test (Lilliefors, 1967) to our data. The null hypothesis of all tests was rejected at the 5% significance level (see normal probability plot examples in Fig. 4a–d).

3. Results

3.1. Packing density

The mean packing density values for mesophilic and psychrophilic proteins show a 0.0057 mean difference between the two groups (Table 2). The calculated *p*-values for the packing densities in these two groups of proteins do not show considerable difference between the homologous pairs. This can probably be explained by Fleming et al.'s finding that packing density is in general similar for conserved domains of homologous proteins within a protein family, while for buried residues in these domains the packing densities are not the same because of distinct secondary structure at different loci (Fleming and Richards, 2000). While in our previous study we have detected only slight changes in the helical content of psychrophilic compared to mesophilic proteins, it can be inferred that packing density may not be significantly different between the psychrophilic–mesophilic homologous protein pairs (Jahandideh et al., 2008). Nevertheless, psychrophilicity rules are completely different from thermophilicity of proteins to expect less compact density for psychrophilic proteins compared to thermophilic proteins (Robinson-Rechavi and Godzik, 2005).

3.2. Cavity in psychrophilic and mesophilic proteins

The calculated number of cavities for each pair shows that despite the increase in the number of cavities for some psychrophilic enzymes like Beta galactosidase (1YQ2), Cold amino peptidase (3CIA) and alpha-amylase (1AQH), this feature does not show significant differences between psychrophilic and mesophilic proteins (*p*-value=0.3). The number of cavities has a direct correlation with molecular weight in thermophilic proteins, which means that increase in molecular weight creates energetically unfavorable packing probability (Hubbard and Argos, 1995). However, in our results we observe an insignificant correlation between molecular

Table 2
Output of statistical tests on physico-chemical parameters[a].

	Parameters						
	ASA	BSA	ROG	MW	HB	PD	Cavity
Psychrophilic mean	0.354 ± 0.056 (SD=0.007)	0.028 ± 0.041 (SD=0.005)	19.474 ± 4.165 (SD=4.165)	109.38 ± 3.381 (SD=3.381)	67.78 ± 5.091 (SD=5.091)	0.717 ± 0.025 (SD=0.025)	0.059 ± 0.025 (SD=0.008)
Mesophilic mean	0.363 ± 0.056 (SD=0.006)	0.036 ± 0.043 (SD=0.006)	19.305 ± 4.702 (SD=4.701)	109.501 ± 4.749 (SD=4.748)	67.513 ± 6.385 (SD=6.385)	0.712 ± 0.012 (SD=0.022)	0.062 ± 0.027 (SD=0.009)
Mean difference	0.0089	0.0082	0.1697	0.1207	0.2666	0.0057	0.0023
t-Test	$p=0.23$	$p=0.07$	$p=0.69$	$p=0.85$	$p=0.66$	$p=0.27$	$p=0.3$

[a] ASA: accessible surface area; BSA: buried surface area; ROG: radius of gyration; MW: molecular weight per residue; HB: hydrogen bond; PD: packing density; SD: standard deviation.

weight and the number of cavities in psychrophilic proteins, suggesting the irrelevance of this rule in psychrophilicity of cold-adapted proteins.

3.3. ASA

Our analysis shows that there is difference of 650.515 Å between mean averages of accessible surface areas in two groups of mesophilic and psychrophilic proteins. As seen in Table 2, *p*-value was calculated for paired *t*-test. This *p*-value does not show a significant difference between mesophilic proteins and psychrophilic ones. One reason might be ASA values are biased, because there is a positive correlation between ASA values of these proteins and their molecular weight as seen in Figs. 2 and 3. To solve this issue, we divided ASA values by their corresponding MW values (ASA/MW). Then, we used paired *t*-test to compare ASA/MW values of mesophilic with those of psychrophilic proteins. The results of this test show that the difference between mean averages of ASA/MW values of these two groups is statistically on the border of significance with a *p*-value of 0.059.

3.4. BSA

To obtain the *p*-value for the BSA, we performed a paired *t*-test. This resulted in a *p*-value that shows an almost but not quite significant difference between two groups of mesophilic and psychrophilic proteins.

3.5. Molecular weight per residue

The molecular weight was calculated per residue for all protein sequences in our database to compare psychrophilic–mesophilic protein pairs. Our results have shown minor differences between single psychrophilic–mesophilic protein pairs, which disagrees with our previous result that suggested significant difference between frequency of single amino acids distribution between psychrophilic and mesophilic homologous pairs (Jahandideh et al., 2007a). To validate this recent observation, we performed *t*-test on this larger dataset obtaining a *p*-value of 0.85 for molecular weight per residue, which means there is no significant difference in molecular weight per residue between psychrophilic compared with mesophilic pairs of proteins (Table 2).

3.6. Radius of gyration (ROG)

The statistical analysis of the ROG of 60 protein structures indicates no significant difference between psychrophilic and mesophilic pairs of proteins. Additionally, *p*-values obtained from the *t*-tests cannot describe any relationship between the ROG of psychrophilic and mesophilic pairs of proteins. This can be explained by the direct correlation between ROG and protein structure compactness (Lobanov et al., 2008). Therefore, as our results have shown no significant difference between ROG of psychrophilic and mesophilic homologs pairs, the obtained *p*-value is in good agreement with literature (Kundu and Roy, 2009). Statistics about the ROG are reported in Table 2.

3.7. Hydrogen bonds (HB)

Previously, we have examined distribution of hydrogen bonds at different positions in psychrophilic and mesophilic proteins and our results only have shown significant difference for mean frequency of H-bond (i, $i+5$) between psychrophilic and mesophilic proteins (Jahandideh et al., 2007a). Here, we examined distribution of total hydrogen bonds and our results have shown no significant differences between mesophilic and psychrophilic

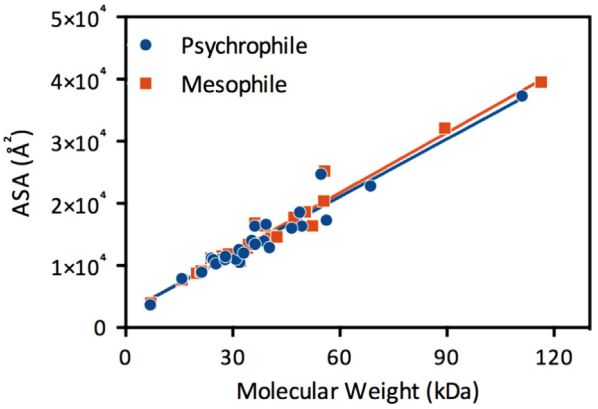

Fig. 3. ASA values are plotted against their corresponding molecular weight. As seen in this figure, there is a positive correlation between ASA values and those of molecular weights.

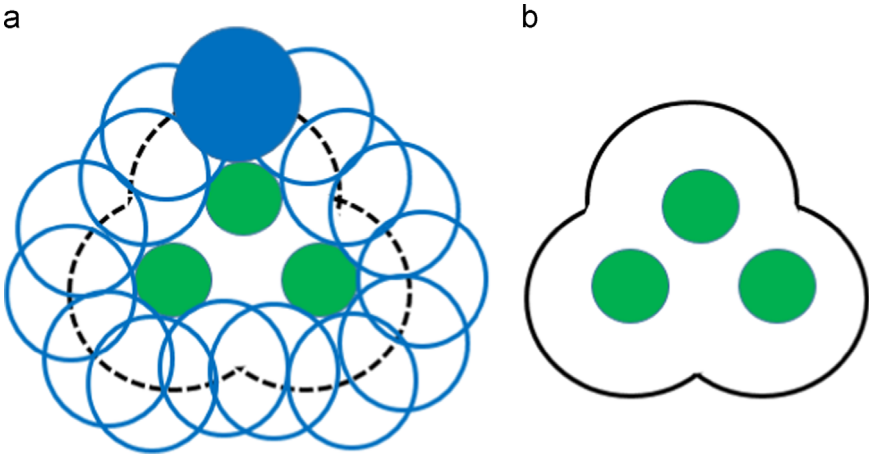

Fig. 2. The accessible surface area of three arbitrary atoms is schematically shown in a two-dimensional plane. The blue circle, green circles and black boundary represent the solvent probe, protein atoms and accessible surface area, respectively. (a) The empty blue circles show the rolling probe sphere and (b) the solid black line represents the calculated accessible surface area. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

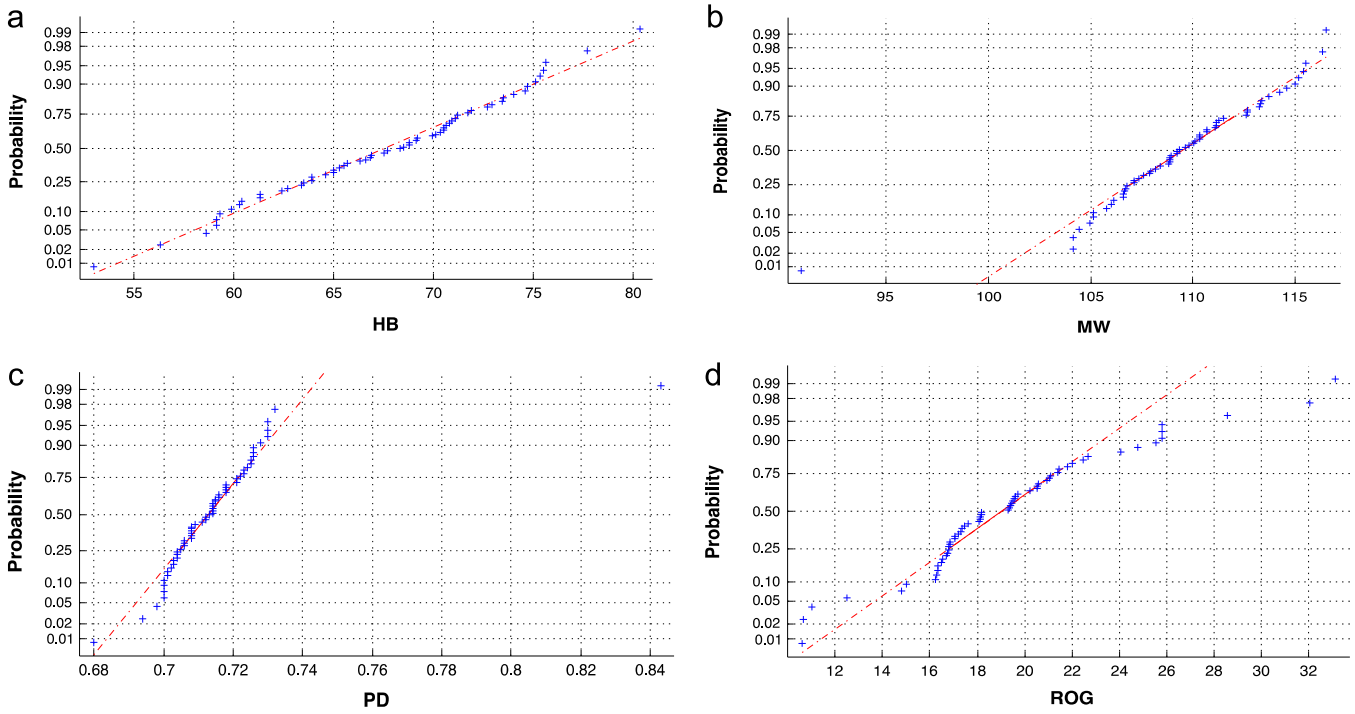

Fig. 4. Normal probability distribution of (a) HB, (b) MW, (c) PD, and (d) ROG.

pairs of proteins (Fig. 4a). Statistics relative to the total number of hydrogen bonds observed in our data set are reported in Table 2.

4. Discussion

It is almost a decade since we initiated studying the structural properties of cold-adapted proteins. We have explored novel rules of phsychrophilicity, e.g. significant difference in total open β-turn content, substitution of amino acids, distribution of amino acids in secondary structures, etc. (Jahandideh et al., 2007a, 2007b, 2008). In this study, we extended our previous database to 30 pairs of structurally defined proteins categorized as psychrophilic and mesophilic to examine a few physicochemical parameters involved in protein cold adaptation.

In this study, we have limited our database to a list of high-resolution structure-defined proteins by X-ray crystallography. However, there is a fact that not all proteins could be identified by X-ray crystallization and NMR spectroscopy, many computational studies have shed light on unknown protein structures (Chou et al., 1997). Integral membrane proteins are among those proteins with a very high difficulty determining structure (Schnell and Chou, 2008). Though NMR has made promising advances in determining tertiary structure of some membrane and transmembrane proteins, but it is also time-consuming and costly (Berardi et al., 2011; Call et al., 2006; Yang et al., 2015). To acquire the structural information in a timely manner, a series of 3D protein structures were developed by means of homology technique (Wang et al., 2007, 2009; Chou, 2004; Chou et al., 2000), and were found very useful for drug development.

Our results confirm that, in general, thermophilicity is correlated with rigidity of a protein, whereas psychrophilicity should be reflected by a more flexible protein structure, which in fact reduces the energetic cost of the conformational changes that are required for interacting with the substrate. In addition, our studies explored a complementary rule for psychrophilicity, which says that among the parameters essential for psychrophilicity, β-turns are mostly local. On the other hand, among the parameters essential for thermostability, higher degree of contact order is mostly global while this is not true for psychrophilic proteins. In conclusion, the evolutionary strategy of adaptation to low temperatures is not merely the inverse of adaptation to high temperatures. Considering that our current knowledge about physico-chemical mechanism of cold adaptation is very limited, our study will assist us in deciphering the adaptive mechanisms that allow psychrophilic organisms to tolerate and survive in cold environments.

Acknowledgments

We want to express our thanks to all members of The First Protein Structural Bioinformatics Online Workshop. This manuscript extracted from part of our experiments on protein structure prediction in this workshop.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.jtbi.2015.07.013.

References

Baldasseroni, F., Pascarella, S., 2009. Subunit interfaces of oligomeric hyperthermophilic enzymes display enhanced compactness. Int. J. Biol. Macromol. 44, 353–360. http://dx.doi.org/10.1016/j.ijbiomac.2009.02.002.

Berardi, M.J., Shih, W.M., Harrison, S.C., Chou, J.J., 2011. Mitochondrial uncoupling protein 2 structure determined by NMR molecular fragment searching. Nature 476, 109–113. http://dx.doi.org/10.1038/nature10257.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E., 2000. The protein data bank. Nucleic Acids Res. 28, 235–242.

Buzzini, P., Branda, E., Goretti, M., Turchetti, B., 2012. Psychrophilic yeasts from worldwide glacial habitats: diversity, adaptation strategies and biotechnological potential. FEMS Microbiol Ecol 82, 217–241. http://dx.doi.org/10.1111/j.1574-6941.2012.01348.x.

Call, M.E., Schnell, J.R., Xu, C., Lutz, R.A., Chou, J.J., Wucherpfennig, K.W., 2006. The structure of the zetazeta transmembrane dimer reveals features essential for its

assembly with the T cell receptor. Cell 127, 355–368. http://dx.doi.org/10.1016/j.cell.2006.08.044.
Cameron, K.A., Hodson, A.J., Osborn, A.M., 2012. Structure and diversity of bacterial, eukaryotic and archaeal communities in glacial cryoconite holes from the Arctic and the Antarctic. FEMS Microbiol Ecol 82, 254–267. http://dx.doi.org/10.1111/j.1574-6941.2011.01277.x.
Cavicchioli, R., 2006. Cold-adapted archaea. Nat Rev Microbiol 4, 331–343. http://dx.doi.org/10.1038/nrmicro1390.
Casanueva, A., Tuffin, M., Cary, C., Cowan, D.A., 2010. Molecular adaptations to psychrophily: the impact of 'omic' technologies. Trends Microbiol. 18, 374–381. http://dx.doi.org/10.1016/j.tim.2010.05.002.
Chen, W., Feng, P.M., Deng, E.Z., 2014. iTIS-PseTNC: a sequence-based predictor for identifying translation initiation site in human genes using pseudo trinucleotide composition. Anal. Biochem. 462, 76–83.
Chen, W., Feng, P.M., Lin, H., Chou, K.C., 2013. iRSpot-PseDNC: identify recombination spots with pseudo dinucleotide composition. Nucleic Acids Res. 41, e68. http://dx.doi.org/10.1093/nar/gks1450.
Chou, K.-C., Jones, D., Heinrikson, R.L., 1997. Prediction of the tertiary structure and substrate binding site of caspase-8. FEBS Lett. 419, 49–54. http://dx.doi.org/10.1016/S0014-5793(97)01246-5.
Chou, K.C., 2004. Structural bioinformatics and its impact to biomedical science. Curr. Med. Chem. 11, 2105–2134.
Chou, K.C., Tomasselli, A.G., Heinrikson, R.L., 2000. Prediction of the tertiary structure of a caspase-9/inhibitor complex. FEBS Lett. 470, 249–256.
Cipolla, A., Delbrassine, F., DaLage, J.L., Feller, G., 2012. Temperature adaptations in psychrophilic, mesophilic and thermophilic chloride-dependent alpha-amylases. Biochimie 94, 1943–1950. http://dx.doi.org/10.1016/j.biochi.2012.05.013.
Criswell, A.R., Bae, E., Stec, B., Konisky, J., Phillips Jr., G.N., 2003. Structures of thermophilic and mesophilic adenylate kinases from the genus Methanococcus. J. Mol. Biol. 330, 1087–1099.
D'Amico, S., Claverie, P., Collins, T., et al., 2002. Molecular basis of cold adaptation. Philos. Trans. R. Soc. B: Biol. Sci. 1423, 917–925. http://dx.doi.org/10.1098/rstb.2002.1105.
Feller, G., 2013. Psychrophilic enzymes: from folding to function and biotechnology. Scientifica 2013, 512840. http://dx.doi.org/10.1155/2013/512840.
Feng, P.M., Chen, W., Lin, H., Chou, K.C., 2013. iHSP-PseRAAAC: identifying the heat shock protein families using pseudo reduced amino acid alphabet composition. Anal. Biochem. 442, 118–125. http://dx.doi.org/10.1016/j.ab.2013.05.024.
Fleming, P.J., Richards, F.M., 2000. Protein packing: dependence on protein size, secondary structure and amino acid composition. J. Mol. Biol. 299, 487–498. http://dx.doi.org/10.1006/jmbi.2000.3750.
Gianese, G., Bossa, F., Pascarella, S., 2002. Comparative structural analysis of psychrophilic and meso- and thermophilic enzymes. Proteins: Structure, Function, and Bioinformatics 47, 236–249. http://dx.doi.org/10.1002/prot.10084.
Goede, A., Preissner, R., Frömmel, C., 1997. Voronoi cell: new method for allocation of space among atoms: elimination of avoidable errors in calculation of atomic volume and density. J. Comput. Chem. 18, 1113–1123, doi:10.1002/(SICI)1096-987X(19970715)18:9<1113::AID-JCC1>3.0.CO;2-U.
Goodchild, A., Saunders, N.F., Ertan, H., Raftery, M., Guilhaus, M., Curmi, P.M., Cavicchioli, R., 2004. A proteomic determination of cold adaptation in the Antarctic archaeon, Methanococcoides burtonii. Mol Microbiol 53, 309–321. http://dx.doi.org/10.1111/j.1365-2958.2004.04130.x.
Grzymski, J.J., Carter, B.J., DeLong, E.F., Feldman, R.A., Ghadiri, A., Murray, A.E., 2006. Comparative genomics of DNA fragments from six Antarctic marine planktonic bacteria. Appl Environ Microbiol 72, 1532–1541. http://dx.doi.org/10.1128/aem.72.2.1532-1541.2006.
Guo, S.H., Deng, E.Z., Xu, L.Q., Ding, H., 2014. iNuc-PseKNC: a sequence-based predictor for predicting nucleosome positioning in genomes with pseudo k-tuple nucleotide composition. Bioinformatics 30, 1522–1529.
Hubbard, S.J., Argos, P., 1995. Detection of internal cavities in globular proteins. Protein Eng. 8, 1011–1015.
Jahandideh, M., Barkooie, S.M., Jahandideh, S., Abdolmaleki, P., Movahedi, M.M., Hoseini, S., Asadabadi, E.B., Jouni, F.J., Karami, Z., Firoozabadi, N.H., 2008. Elucidating the protein cold-adaptation: investigation of the parameters enhancing protein psychrophilicity. J. Theor. Biol. 255, 113–118. http://dx.doi.org/10.1016/j.jtbi.2008.07.034.
Jahandideh, S., Abdolmaleki, P., Jahandideh, M., Barzegari Asadabadi, E., 2007a. Sequence and structural parameters enhancing adaptation of proteins to low temperatures. J. Theor. Biol. 246, 159–166. http://dx.doi.org/10.1016/j.jtbi.2006.12.008.
Jahandideh, S., Barzegari Asadabadi, E., Abdolmaleki, P., Jahandideh, M., Hoseini, S., 2007b. Protein psychrophilicity: role of residual structural properties in adaptation of proteins to low temperatures. J. Theor. Biol. 248, 721–726. http://dx.doi.org/10.1016/j.jtbi.2007.06.019.
Janin, J., Miller, S., Chothia, C., 1988. Surface, subunit interfaces and interior of oligomeric proteins. J. Mol. Biol. 204, 155–164.
Kandaswamy, K.K., Chou, K.C., Martinetz, T., Moller, S., Suganthan, P.N., Sridharan, S., Pugalenthi, G., 2011. AFP-Pred: a random forest approach for predicting antifreeze proteins from sequence-derived properties. J. Theor. Biol. 270, 56–62. http://dx.doi.org/10.1016/j.jtbi.2010.10.037.
Krissinel, E., Henrick, K., 2007. Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. 372, 774–797. http://dx.doi.org/10.1016/j.jmb.2007.05.022.
Kundu, S., Roy, D., 2009. Comparative structural studies of psychrophilic and mesophilic protein homologues by molecular dynamics simulation. J Mol Graph Model 27, 871–880. http://dx.doi.org/10.1016/j.jmgm.2009.01.004.
Lauro, F.M., McDougald, D., Thomas, T., Williams, T.J., Egan, S., Rice, S., DeMaere, M. Z., Ting, L., Ertan, H., Johnson, J., Ferriera, S., Lapidus, A., Anderson, I., Kyrpides, N., Munk, A.C., Detter, C., Han, C.S., Brown, M.V., Robb, F.T., Kjelleberg, S., Cavicchioli, R., 2009. The genomic basis of trophic strategy in marine bacteria. Proc Natl Acad Sci U S A 106, 15527–15533. http://dx.doi.org/10.1073/pnas.0903507106.
Lee, B., Richards, F.M., 1971. The interpretation of protein structures: estimation of static accessibility. J. Mol. Biol. 55, 379–400.
Lilliefors, H.W., 1967. On the Kolmogorov–Smirnov test for normality with mean and variance unknown. J. Am. Stat. Assoc. 62, 399–402.
Lin, H., Deng, E.Z., Ding, H., Chen, W., 2014. iPro54-PseKNC: a sequence-based predictor for identifying sigma-54 promoters in prokaryote with pseudo k-tuple nucleotide composition. Nucleic Acids Res. 42, 12961–12972.
Liu, B., Fang, L., Liu, F., Wang, X., Chou, K.C., 2015a. iMiRNA-PseDPC: microRNA precursor identification with a pseudo distance-pair composition approach. J Biomol Struct Dyn, 1–13. http://dx.doi.org/10.1080/07391102.2015.1014422.
Liu, Z., Xiao, X., Qiu, W.R., 2015b. iDNA-Methyl: identifying DNA methylation sites via pseudo trinucleotide composition. Anal. Biochem. 474, 69–77.
Lobanov, M.Y., Bogatyreva, N.S., Galzitskaya, O.V., 2008. Radius of gyration as an indicator of protein structure compactness. Mol. Biol. 42, 623–628. http://dx.doi.org/10.1134/S0026893308040195.
Marsh, J.A., 2013. Buried and accessible surface area control intrinsic protein flexibility. J. Mol. Biol. 425, 3250–3263. http://dx.doi.org/10.1016/j.jmb.2013.06.019.
McDonald, I.K., Thornton, J.M., 1994. Satisfying hydrogen bonding potential in proteins. J. Mol. Biol. 238, 777–793. http://dx.doi.org/10.1006/jmbi.1994.1334.
Methe, B.A., Nelson, K.E., Deming, J.W., Momen, B., Melamud, E., Zhang, X., Moult, J., Madupu, R., Nelson, W.C., Dodson, R.J., Brinkac, L.M., Daugherty, S.C., Durkin, A. S., DeBoy, R.T., Kolonay, J.F., Sullivan, S.A., Zhou, L., Davidsen, T.M., Wu, M., Huston, A.L., Lewis, M., Weaver, B., Weidman, J.F., Khouri, H., Utterback, T.R., Feldblyum, T.V., Fraser, C.M., 2005. The psychrophilic lifestyle as revealed by the genome sequence of Colwellia psychrerythraea 34H through genomic and proteomic analyses. Proc. Natl. Acad. Sci. USA 102, 10913–10918. http://dx.doi.org/10.1073/pnas.0504766102.
Morita, R.Y., 1975. Psychrophilic bacteria. Bacteriol. Rev. 39, 144–167.
Ooi, T., Oobatake, M., Némethy, G., Scheraga, H.A., 1987. Accessible surface areas as a measure of the thermodynamic parameters of hydration of peptides. Proc. Natl. Acad. Sci. USA 10, 3086–3090.
Pack, S.P., Yoo, Y.J., 2005. Packing-based difference of structural features between thermophilic and mesophilic proteins. Int. J. Biol. Macromol. 35, 169–174. http://dx.doi.org/10.1016/j.ijbiomac.2005.01.007.
Paredes, D.I., Watters, K., Pitman, D.J., Bystroff, C., Dordick, J.S., 2011. Comparative void-volume analysis of psychrophilic and mesophilic enzymes: structural bioinformatics of psychrophilic enzymes reveals sources of core flexibility. BMC Struct. Biol. 11, 42. http://dx.doi.org/10.1186/1472-6807-11-42.
Qiu, W.R., Xiao, X., 2014. iRSpot-TNCPseAAC: identify recombination spots with trinucleotide composition and pseudo amino acid components. Int. J. Mol. Sci. 15, 1746–1766.
Richmond, T.J., 1984. Solvent accessible surface area and excluded volume in proteins. Analytical equations for overlapping spheres and implications for the hydrophobic effect. J. Mol. Biol. 178, 63–89.
Robinson-Rechavi, M., Godzik, A., 2005. Structural genomics of thermotoga maritima proteins shows that contact order is a major determinant of protein thermostability. Structure 13, 857–860. http://dx.doi.org/10.1016/j.str.2005.03.011.
Saunders, N.F., Thomas, T., Curmi, P.M., Mattick, J.S., Kuczek, E., Slade, R., Davis, J., Franzmann, P.D., Boone, D., Rusterholtz, K., Feldman, R., Gates, C., Bench, S., Sowers, K., Kadner, K., Aerts, A., Dehal, P., Detter, C., Glavina, T., Lucas, S., Richardson, P., Larimer, F., Hauser, L., Land, M., Cavicchioli, R., 2003. Mechanisms of thermal adaptation revealed from the genomes of the Antarctic Archaea Methanogenium frigidum and Methanococcoides burtonii. Genome Res. 13, 1580–1588. http://dx.doi.org/10.1101/gr.1180903.
Schnell, J.R., Chou, J.J., 2008. Structure and mechanism of the M2 proton channel of influenza A virus. Nature 451, 591–595. http://dx.doi.org/10.1038/nature06531.
Shrake, A., Rupley, J.A., 1973. Environment and exposure to solvent of protein atoms. Lysozyme and insulin. J. Mol. Biol. 79, 351–371.
Simon, I., 1971. Determination of small alterations in the radius of gyration by small-angle X-ray scattering. J. Appl. Crystallogr. 4, 317–318.
Srinivas, G., Ochije, H., Mohan, R., 2014. Biomolecules in binary solvents: computer simulation study of lysozyme protein in ethanol–water mixed solvent environment. JSM Nanotechnol. Nanomed. 2, 1029.
Stellwagen, E., Wilgus, H., 1978. Relationship of protein thermostability to accessible surface area. Nature 275, 342–343.
Wang, J.F., Wei, D.Q., Li, L., Zheng, S.Y., Li, Y.X., Chou, K.C., 2007. 3D structure modeling of cytochrome P450 2C19 and its implication for personalized drug design. Biochem. Biophys. Res. Commun. 355, 513–519. http://dx.doi.org/10.1016/j.bbrc.2007.01.185.
Wang, S.Q., Du, Q.S., Huang, R.B., Zhang, D.W., Chou, K.C., 2009. Insights from investigating the interaction of oseltamivir (Tamiflu) with neuraminidase of the 2009 H1N1 swine flu virus. Biochem. Biophys. Res. Commun. 386, 432–436. http://dx.doi.org/10.1016/j.bbrc.2009.06.016.

Wu, Z.C., Xiao, X., 2010. 2D-MH: a web-server for generating graphic representation of protein sequences based on the physicochemical properties of their constituent amino acids. J. Theor. Biol. 267, 29–34.
Xiao, X., Wang, P., 2012. iNR-PhysChem: a sequence-based predictor for identifying nuclear receptors and their subfamilies via physical–hemical property matrix. PLoS One 7, e30869.
Xiao, X., Lin, W.Z., 2013. Recent advances in predicting protein classification and their applications to drug development. Curr. Top. Med. Chem. 13, 1622–1635.
Xiao, X., Min, J.L., Wang, P., 2013a. Predict drug–protein interaction in cellular networking. Curr. Top. Med. Chem. 13, 1707–1712.
Xiao, X., Wang, P., 2013. Recent progresses in identifying nuclear receptors and their families. Curr. Top. Med. Chem. 13, 1192–1200.
Xiao, X., Wang, P., Lin, W.Z., 2013b. iAMP-2L: a two-level multi-label classifier for identifying antimicrobial peptides and their functional types. Anal. Biochem. 436, 168–177.
Yang, J.J., Chen, Y., Lu, G., Tzoneva, G., Ma, X., Wu, T., Li, W., Lu, H., Ding, L., Liang, H., Huang, X., Yang, M., Jin, L., Kang, H., Chen, S., Du, A., Shen, S., Ding, J., Chen, H., Chen, J., von Stackelberg, A., Gu, L., Zhang, J., Ferrando, A., Tang, J., Wang, S., Zhou, B.B., 2015. Nat. Med., http://dx.doi.org/10.1038/nm.3840.