

# On the Role of Water in the Protein Activity

L. Degrève, G.H. Brancaloni, C.A. Fuzo, M.R. Lourenzoni,  
F.M. Mazzé, A.M. Namba, and D.S. Vieira

*Grupo de Simulação Molecular, Departamento de Química,  
Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo  
14040-350 Ribeirão Preto (S.P.), Brazil*

Received on 05 October, 2003.

The role of the supporting medium of water molecules in some protein activities is examined under different aspects as in the cases of a monomeric peptide, the *basic fibroblast growth factor*, of a dimeric peptide, the *human neutrophil peptide 3*, of a peptide that acts in non-aqueous environment, the *gramicidin A* dimer, of a water molecule present in the binding of a co-factor in a *phospholipase* peptide, and under the general point of view of the hydrophilic/hydrophobic properties described by a hydrophathy scale. These examples illustrate the importance of water in the hydrogen bond formation that is, of main importance in keeping the peptide structures that cannot be defined without the water contributions. The conclusions confirm that living systems are like they are because water is an outstanding and abundant molecule present everywhere in living matter.

## 1 Introduction

Almost all mechanisms occurring in the cells depend on proteins [1]-[6] that constitute most of biological macromolecules providing large variety of functions[2]-[4]. Their importance can be emphasized by noting that the genetic information is fundamentally expressed as protein molecules[5]-[7]. Specific DNA segments contain the genetic information on the peptide amino acids sequence. Thousands of different proteins in the cells, precisely codified by the genes, realize specific functions. The individuality of the proteins is directly related with their three-dimensional structure that provides the ideal conditions for realizing correctly their functions[8]. However, in spite of the importance of the proteins in the cellular life, the role of the supporting medium, in particular of the solvent molecules, *i.e.* the water molecules, is equally important because they constitute the most part of the cell, about 70%, and because the contribution of the water molecules is essential in the peptide activities[9]. The water molecules[10, 11], beyond the small mass, present high multipolar moments that contribute to the formation of hydrogen bonds, HB. Consequently, the water molecules perform an essential structural role in the organization and activity of the biological medium. Many of the protein activities depend on the protein stability, on associations with other proteins or ligands while the catalytic activity depend on the structure, on thermodynamic and dynamic properties, properties that are deeply influenced by the solvent.

All the most important biological molecules like pep-

tides, saccharides, nucleic acids have the common feature that they contain hydrogen-bonding functional groups[8]. The hydrogen bonds have some interesting features like: (i) low enthalpy formation, about 20% of the chemical bond enthalpy; (ii) total unspecificity; (iii) the association, or dissociation, of molecules through HB are fast enough to permit to check their formation and to correct misformations. The large number of HB, that are generally present between water and biological materials, result in the high specificity of the HB networks which include intra and intermolecular bonds.

In the present paper, the role of the water molecules in the intermolecular HB formations in biological media will be described by means of some typical examples. In the resulting analysis, the consequences of the water molecules participation on the peptide structures will be investigated as well. The direct role of the water molecules in the definition of protein structures will be examined in a first example as the case of a monomeric peptide, the *basic fibroblast growth factor*, in a second example as the case of a dimeric peptide, the *human neutrophil peptide 3*. The third analysis will focus a peptide that acts in non-aqueous environment, the *gramicidin A* dimer and, as a fourth example, the role of a water molecule present in the binding of the co-factor will be analyzed in the *phospholipase A2* case. The final example will be the analysis of a hydrophathy scale that is a scale helpful in the determination of three-dimensional structures of the peptides since it makes a relationship between the affinity of the individual amino acid residues and their location in a proteic structure.

## 2 Material

### 2.1 The basic fibroblast growth factor

The function of the growth factors peptides is to induce their specific target cells to grow and/or to make a differentiation. The first step in their activity is to bind to heparin or to the heparan sulfate chains of proteoglycans located on the cell-surface receptor[2]-[4]. The *basic fibroblast growth factor* (*bFGF* or *FGF-2*) is a mitogenic, neurotrophic, and angiogenic polypeptide that is a member of a protein family containing until the present nine known different species (constituting the *FGFs* family) that interact with three, related but distinct, receptors[13, 14]. Several members of the *FGFs* family are oncogenes: the angiogenic properties of *bFGF* suggest an involvement in tumor growth and cancer. The *basic fibroblast growth factor* also has wound healing properties that made it an attractive candidate as a therapeutic drug. Even though the three-dimensional structure of *bFGF* has been recently elucidated relatively little is known about its structure-function relationship. Two regions in the primary structure of the 155 amino acid *bFGF* have been proposed to be involved in the receptor-binding and mitogenic activity of this factor. These regions correspond to residues 33-77 and 115-124[6, 12, 16].

The secondary structure of *bFGF* comprises twelve antiparallel  $\beta$  strands arranged in a pattern with approximate threefold internal symmetry[6, 12, 16]. The strands are numbered sequentially from the amino terminus. The  $\beta$  strands 1, 4, 5, 8, 9 and 12 form a six stranded antiparallel  $\beta$  barrel that is closed at one end by  $\beta$  sheet interactions involving strands 2, 3, 6, 7, 10 and 11. The three-dimensional structure of the *basic fibroblast growth factor* is displayed in Fig. 1.

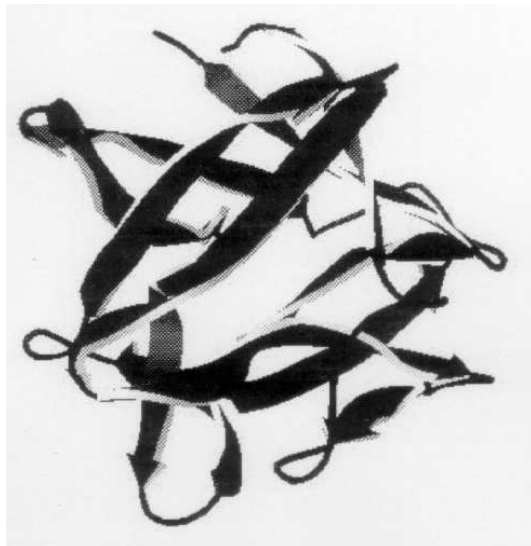


Figure 1. The tertiary structure of the *basic fibroblast growth factor*.

### 2.2 The defensin HNP-3

Antibacterial cationic peptides are important components of the innate defenses of all species of life[17]. Most of them interact with the bacterial membranes by disrupting the order of the phospholipid bilayer, causing loss of membrane integrity [18]-[20]. Two main mechanisms have been suggested for peptide permeation of the bacterial membrane: (i) the barrel-stave mechanism, where bundles of peptides form transmembrane pores through the bacterial membrane[21] and (ii) the carpet-like mechanism, where membrane destruction/solubilization occurs via parallel binding of the peptides to the bacterial membrane, covering the membrane in a carpet like manner[22].

The human neutrophil peptide 3, *HNP-3*, is a member of the  $\alpha$ -defensin family, one of the most common classes of cationic antimicrobial peptides[21, 23]-[26]. The primary structure of *HNP-3* is displayed in the Fig. 2. The common structural feature of the  $\alpha$ -defensin family is a hydrogen-bonded pair of antiparallel  $\beta$ -strands (strands 2 and 3) connected by a short turn to form a  $\beta$ -hairpin. The crystal structure of *HNP-3*[26] shows a dimeric structure formed by two monomers connected by their  $\beta$ -hairpin regions through HB. The resulting quaternary structure is a six-stranded  $\beta$ -sheet (Fig. 3) stabilized by HB, six disulfide bridges and hydrophobic contacts. Moreover, the crystallographic analysis shows ordered internal water molecules stabilizing the peptide structure[26]. The tertiary structures of the two monomers, monomer 1 and monomer 2, are slightly different.

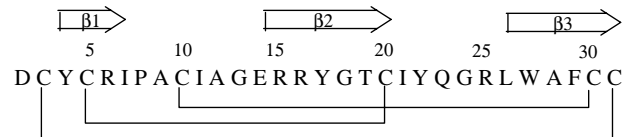


Figure 2. Primary structure of the *HNP-3*. The disulfide bridges are represented by the solid lines.

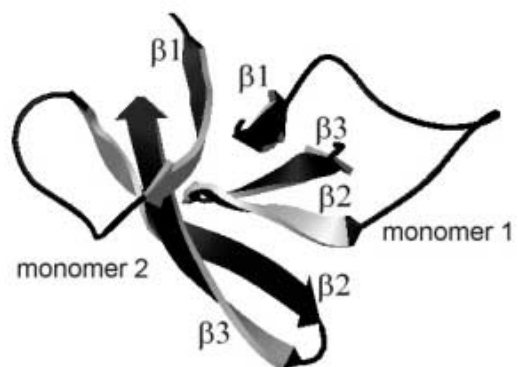


Figure 3. The *HNP-3* quaternary structure.

### 2.3 The gramicidin A

The dimer *gramicidin A* (GA) is a 30 amino acid residues peptide with antibiotic properties that is produced by *Bacillus brevis*. GA is active against Gram-positive bacterial membranes[27, 28] by forming one of the best-characterized ion channels. The GA channel increases the membrane permeability to water molecules and to monovalent ions[2]-[4]. It has been used as a model for ion channels in numerous experimental and theoretical studies. It has been accepted that GA is a N-terminal to N-terminal right handed  $\beta$ -helix dimer during the channel formation[4, 29, 30]. GA is constituted by an amino acid residues sequence where D and L configurations are alternated. The N-terminus is formylated and the C-terminus is bound to an ethanolamine group. The secondary structure is depicted in the Fig. 4. The GA biological activity is intimately related to its structure so that it may be influenced by different factors such as the nature of the solvent[8, 31]-[34]. Consequently the importance of the water molecules in the maintenance of the structure of the GA in the active form is a main object of study.



Figure 4. The secondary structure of *gramicidin A*.

### 2.4 Phospholipases A2

*Phospholipases A2* (PLA2 EC 3.1.1.4) catalyze the hydrolysis of the sn-2 acyl bonds of sn-3 phospholipids. Snake venoms constitute a rich source of PLA2 with several structural and functional diversity[35]. The classification of venom PLA2s into two classes is based on the basis of primary structure[36]. Classes I/II snake venom PLA2s display several pharmacological properties[37]. The structural basis of the catalytic function involves the highly conserved active site residues His48, Asp49, Tyr52 and Asp99, in which the essential  $\text{Ca}^{2+}$  co-factor is bound to Asp49 and to carbonyl main chain oxygen atoms of the aptly named calcium binding loop. A schematic diagram of the interaction between *phospholipase A2* and a phospholipid is observed in Fig. 5[38]. The mechanism of action of *phospholipase A2* involves a His48 and Asp99 catalytic diad that activates one water molecule (into the dashed circle in Fig. 5) for nucleophilic attack on the ester while  $\text{Ca}^{2+}$  stabilizes the oxyanion transition state[38]. A sub-family of catalytically inactive PLA2s has been characterized with Asp49 substituted

by Lys49. The presence of the  $\text{NH}_3^+$  group of Lys49 makes impossible the binding of the  $\text{Ca}^{2+}$  co-factor resulting in the lack of catalytic activity. However, despite the lack of catalytic activity, these Lys49-PLA2 homologues retain a  $\text{Ca}^{2+}$  independent membrane damaging activity. In spite of the lack of catalytic activity be related in many works, some papers refute this fact and report a residual catalytic activity[39, 40]. Furthermore, as such as in the case of other snake venoms, PLA2s isolated from snake venom, Lys49-PLA2s, also processes combined myotoxic and cytolytic pharmacological activities.

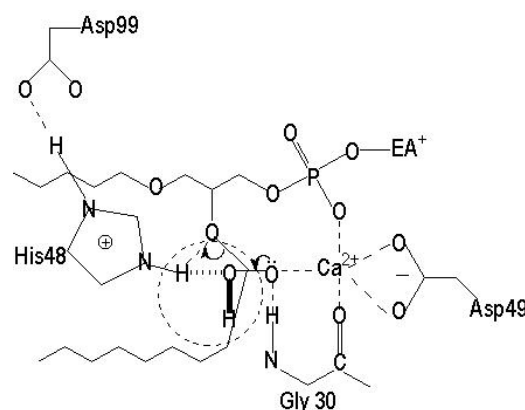


Figure 5. Schematic diagram of a productive interaction between *phospholipase A2* and a phosphatidylethanolamine.

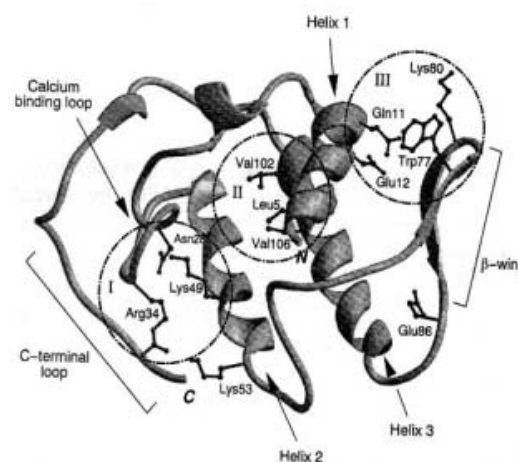


Figure 6. Ribbon representation of the Lys49-PLA2, in which are indicated the different secondary structures and the catalytic site (I).

The general structure of the Lys49-PLA2s monomer is showed in the Fig. 6. The catalytic site, formed by the four residues His48, Lys49, Tyr52 and Asp99, is sheltered by two antiparallel  $\alpha$ -helices. The Lys49-PLA2s side chain N atom of the Lys49 residue is located exactly at the same position occupied by the  $\text{Ca}^{2+}$  ion in the active site of the Asp49-PLA2s[35]. Studies of site-directed autogenesis realized with porcine pancreatic PLA2, in which Asp49 is substituted by Lys49, have suggested that the Lys49 side chain N atom

sterically hinders the binding of the essential co-factor  $\text{Ca}^{2+}$  [35]. The C-terminal region and a single double-stranded  $\beta$ -sheet are linked by disulphide bridges. Several studies have identified amino acid clusters located in the active site and lipid substrate binding regions[35] however the mechanisms of myotoxic and cytolytic activities of Lys49PLA2s are still unknown. *Myotoxin II*, (BaspMT-II), isolated from *Bothrops Asper*, is a basic dimeric Lys49-PLA2s. Molecular simulation was employed to study the hydration structure and the structural changes in monomer of the BaspMT-II.

## 2.5 An amphipathy scale

Considering the fundamental importance in the knowledge of the structure of proteins to understand their functions, notable efforts have been dedicated to predict and to elucidate the three-dimensional shape of amino acids sequences. Some works were directed to the analysis of the protein's hydrophobicity using amphipathy scales that show the relative hydrophobicity of amino acids. In these researches, different pathways were pursued based on: the transfer free energies[41]-[43], the accessible amino acid surfaces in proteins[44], statistical techniques using soluble proteins[45], the interactions between amino acids[46], the modification of existent experience based data scales[47], statistical methods using membrane proteins[48], etc. In some of the above cited cases, different procedures were used: for example, the determination of the free energy of transfer was determined using water and different solvents. The transfer free energy was obtained from the equilibrium between water solutions and: ethanol or dioxane[41], *N*-methylacetamine[42] and *n*-octanol or palmitoyllecithin phosphocholine solutions[43].

It can be noted from the analysis of the amphipathy scales that the hydrophobicities of the amino acids are lacking of uniformity because the residue hydrophobicity is too much correlated with the different kind of methods used in its determination. In the light of the general observation that the interior of soluble proteins is predominantly composed by hydrophobic amino acids, while the hydrophilic side chains are preferentially located on the external surface of the proteins where they are free to interact with the solvent molecules, the present example shows how it is possible to elaborate a new amphipathy scale that determines *in situ* the hydrophobicity of amino acids in water by molecular simulation without using any arbitrary equilibrium. It is highly probable that molecular simulation is to be considered as the good method to weigh correctly the hydrophobicity of the amino acids because this technique is able "to reproduce" the environment protein-water as so as characteristics that depend on the physical and chemical properties of water[49, 50, 51, 52].

## 3 Method

All the results were obtained from molecular simulation using the molecular dynamics method[53]-[55]. The simulated systems consist of the solute molecule(s) immersed in water molecules or in a mixed medium water-carbon tetrachloride in the GA case. The solvent concentrations were adjusted to match, in the pure solvent regions, the experimental solvent density. In the aqueous phases, ions were introduced to neutralize electrically the system. The periodic boundary conditions and minimum image convention were applied[53]-[55]. The Gromos96 force field was used to model all the molecules and interactions[56] including the spc/e model for water molecules. The bond lengths were controlled using the SHAKE constraint algorithm[55] as so as to maintain the rigidity of the solvent molecules. Experimental X-ray or NMR data were used as initial guess for the peptides coordinates excluding in the amphipathy calculations that start with random configurations. The molecular dynamic simulations were conducted at 298 K during 2.0 to 4.0 ns after relaxing the systems. The equations of motion were integrated using the Verlet algorithm with a 2.0 fs time step. A cut-off was applied at 1.0 nm and a generalized Poisson-Boltzmann cavity field method was used to take into account the long-range interactions corrections[57].

The intermolecular HB were characterized by two criteria. The first one is based on the radial distribution function profiles, *rdf*. The *rdf*  $g_{A,O}(r)$  and  $g_{B,H}(r)$  must present a clear definition of a peak in an appropriate region defined by the nature of A, or B, if A and B are respectively a positively and a negatively charged solute atom. The water oxygen and hydrogen atoms are labeled O (Ow) and H (Hw) respectively. Simultaneously, a second criterion was applied to define the intermolecular HB: the distribution of the interaction energies between the solute atoms and the solvent molecules, the pair energy distributions, *P(E)*, must present a peak, or a shoulder, in the attractive region[58, 59]. The number of HB,  $n_{HB}$ , on the chosen atoms was calculated from the usual integration of the *rdf* up to its first minimum and by the number of water molecules that generates the *P(E)* peak exclusively if it is located in the attractive energy region. In some cases, when only a shoulder in *P(E)* is observed in place of a peak,  $n_{HB}$  was calculated taking into account all the solvent molecules that have presented interaction energies with the selected solute atom, lower, or equal, to -5.0 kcal/mol. The values of  $n_{HB}$  obtained from both procedures are highly consistent. The mean energy of the intermolecular HB,  $E_{HB}$ , is also consistently obtained from *P(E)*. The intramolecular HB were detected by two criteria: maximum of the radial distribution function (*rdf*) at a maximum hydrogen-acceptor distance of 0.235nm and an occurrence fraction, *Fr*, larger than 0.1. The structural stability of the protein was monitored by means of the root mean square deviation (*rmsd*).

## 4 Results and discussion

### 4.1 The basic fibroblast growth factor

The initial structure used in the simulation is the structure obtained by X-ray crystallographic diffraction (pdb code: 1BFF[60]). As a first analysis, the structural stability of *bFGF* was investigated from the *rmsd* data that are plotted in Fig. 7. The *rmsd* calculated using the X-ray structure as the reference. These data show that the simulated structures are very similar to the experimental one confirming that the simulation data are reliable for an investigation of the *bFGF* structure.

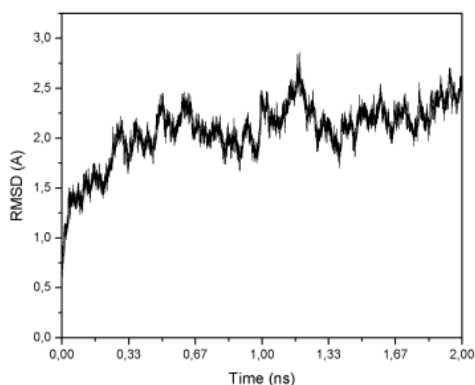


Figure 7. The root mean square deviations for the *basic fibroblast growth factor*.

The backbone hydration was analysed by considering the carbonyl CO and amide NH atoms that are hydrogen bonded to water molecules. Similar conclusions are given by the analysis of the rdfs and P(E) distributions in the case of intermolecular HB formation. The  $g_{CO,H}(r)$  are characterized by a first peak with a maximum located at 1.80-1.90 Å, while the  $g_{NH,O}(r)$  feature first peaks with maxima at 1.75-2.00 Å. The CO P(E) distributions present a peak in the intermolecular attractive region with maxima located around -6 kcal/mol. However, in some cases, the CO-water pair energies give rise to distributions with a shoulder in the intermolecular attractive region instead of a peak. The results for  $n_{HB}$ , shown in Figs. 8 and 9, indicate that 52% of the CO and and 30% of the NH atoms of the *bFGF* forms intermolecular HB.

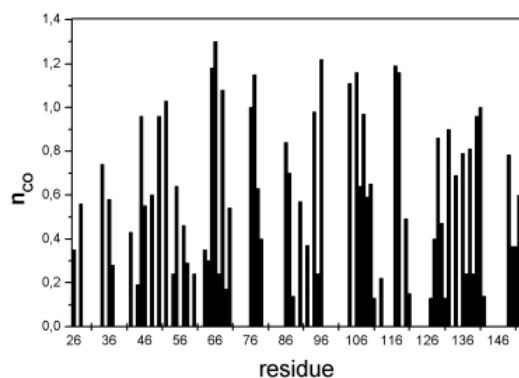


Figure 8. The HB of the *basic fibroblast growth factor* CO atoms.

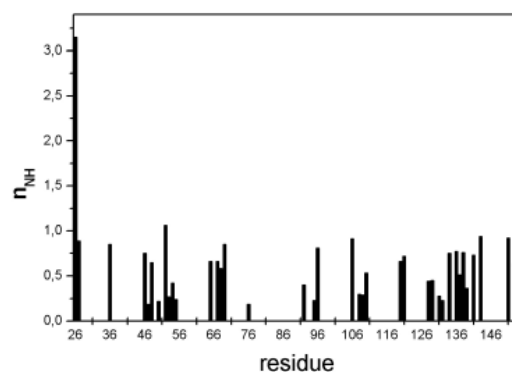


Figure 9. The HB of the *basic fibroblast growth factor* NH atoms.

The degree of exposition to the solvent of the *bFGF* side chains is appropriately described focusing on the interaction of polar or charged groups with water molecules. As such as it was found in the backbone hydration, each functional group exhibits a solvation pattern characterized by rdfs with similar profiles and peak positions. However, these rdfs differ in intensity. The hydration of the lateral chains was analyzed considering only the polar lateral chains, charged or not, since no intermolecular HB were found in the apolar side chains.

The  $g_{HZ,O}(r)$  for the lysine residues HZ atoms exhibit similar profiles featured by peaks with maxima at 1.95 - 2.00 Å. The  $g_{HE,O}(r)$  and  $g_{HH,O}(r)$  functions HE and HH arginine atoms present distinct profiles indicating that the hydration structures around the NH and NH<sub>2</sub> groups are different. In the cases of the arginine residues, the HE-Ow rdfs present similar first peaks located at 2.05 Å. In the HH-Ow rdfs peaks were observed with maxima at 1.90 Å. The pair energy distributions for the interactions between water molecules and the HZ, and HE/HH atoms of the lysines and arginines, respectively, reveal that these interaction energies span from -6.5 to 8.75 kcal/mol. The OE1-Hw, OE2-Hw, OD1-Hw, OD2-Hw rdfs for the aspartic and glutamic acid carboxylate oxygen atoms present a first peak at 2.00 Å and 2.05 Å, respectively. The corresponding interaction energies give rise to distributions that span from -10 to -9 kcal/mol, characterized by peaks in the region of attractive interactions. The carboxyl groups of the aspartic and glutamic acids are similarly solvated as such as the solvations of their amino groups are similar.

The hydration structures of the hydroxyl groups of the serine, tyrosine, threonine residues exhibit similar patterns. Analysis of the rdfs OH-Hw presents peaks at 1.75 Å. The HH-Ow rdfs present a striking sharp first peak at 1.85 - 2.05 Å indicating a strong hydrogen bond. The pair energy distributions for the OH atoms reveal that their hydrogen bonds with water molecules at energies in the range from -9.5 kcal/mol. The pair energy distributions for the HH atoms occur through energies in the range from -8.3 to -13.5 kcal/mol.

The number  $n_{HB}$  of the side chain atoms indicate that some side chains atoms are involved in both intramolecular

and intermolecular hydrogens bonds. In these cases, the hydration numbers are smaller than the hydration numbers of the atoms that are only hydrogen bonded to water, Figs. 8 and 9.

The reasons of the structural stabilization of *bFGF* are clear from the detailed analysis of the set of intermolecular hydrogen binding. The structural stabilization promoted by the HB intramolecular were complemented by the interactions with solvent. The nucleus of *bFGF* is practically inaccessible to the solvent while the external region of this globular protein is in immediate contact with the solvent since 76.7% of the residues of *bFGF* are solvated. These residues can be divided in a group of 36.6% that makes only intermolecular HB while the other 40,1% are involved in both intra and intermolecular HB. The external surface of the protein is rich in hydrophilic and charged amino acids, in particular lysine, arginine, glutamic acid and aspartic acid residues. Consequently the intense hydration of the residues promotes the existence of a peptide/solvent interface where the intermolecular HB insert the surface structure of the protein in the water HB network resulting in a good stabilization of the protein/water contacts. The amide and carboxyl terminal group are ionized at physiological pH and strongly solvated. Hydroxyl groups are present in many solvated side chains contributing by means of intermolecular HB formed to the more attractive intermolecular energies. These energies are so attractive that they surpass the energies of the hydrogen bonds between water molecules in pure water phase [10, 11]. The hydration of the hydroxyl groups is therefore a decisive factor in the stabilization of *bFGF* in aqueous solution.

## 4.2 The defensin HNP-3

The structural organization due to the water near the peptide interface together with intramolecular HB were found to stabilize the crystal structure. The intramolecular HB are: OD1(Asp2)-NH(Cys3), OE1(Glu14)-HE(Arg6), OE2(Glu14)-HH2(Arg6), OE1(Glu14)-NH(Ile11), OH(Tyr17)-HH2(Arg15), OH(Tyr17)-HE(Arg15), CO(Gly18)-HE(Arg16), OG1(Thr19)-HH2(Arg16) and O(Cys31)-NH(Arg15).

The polar amino acids side chains are directed to the polar solvent in aqueous solution differently of the intramolecular HB found in the crystal structure (Fig. 10, crystal structure, and Fig. 11, aqueous solution structure). In solution, the donor and acceptor atoms are involved in intermolecular HB with water molecules disrupting the HB found in solid state. The  $n_{HB}$  values of the polar side chain groups in the monomer 1, Figs. 12-15, indicate that each peptide atom exhibits its own hydration shell. Similar results are observed with the monomer 2. The HNP-3 aqueous solution structure shows that the six Arg residues (Arg6, Arg15 and Arg16 of the both monomers) form an equatorial ring around the dimer, while the apolar side chains of Tyr17, Tyr22, Trp27, Phe29, Cys5 and Cys20 are interacting with each other, repelling water molecules, resulting in a hydrophobic core (Fig. 11). The hydrophobic mini-channel (Fig. 16), that

completely crosses the dimer, does not contain conserved or structural water molecules. However, the presence of the hydration water molecules at the beginning (Asp2 and Tyr4) and at the end (Tyr17 and Tyr22) of the mini-channel preserves the dimer stability. The existence of a hydrophobic mini-channel stabilized by hydration molecules at both ends, in its turn, stabilizes the overall quaternary structure of the *defensin*. Such quaternary structure probably can be easily destroyed when the exposed polar groups begin to interact with membrane surfaces damaging the bilayers structures.

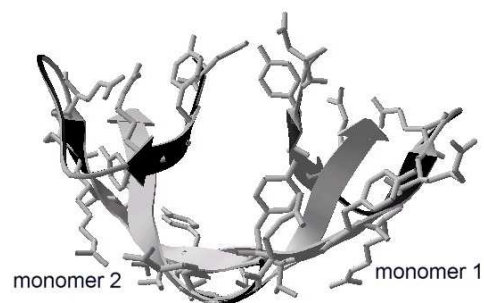


Figure 10. The HNP-3 crystal structure. The polar amino acids are displayed.

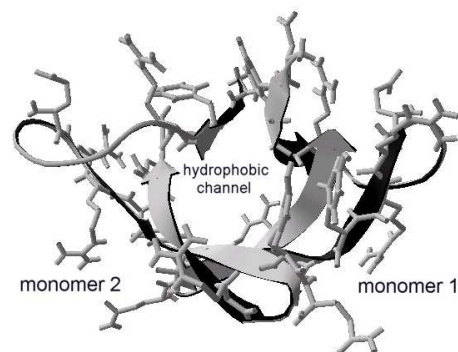


Figure 11. The structure of HNP-3 in aqueous solution with the display of the polar amino acids. The polar side chains, involved in intramolecular HB in the crystal structure, are directed to the polar solvent in the aqueous medium, where they form an equatorial ring around the dimer.

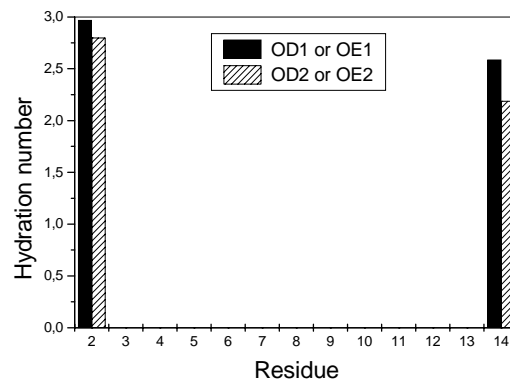


Figure 12. The HB of the aspartic (OD1 and OD2) and glutamic (OE1 and OE2) acid carboxylate oxygen in the HNP-3 monomer 1.

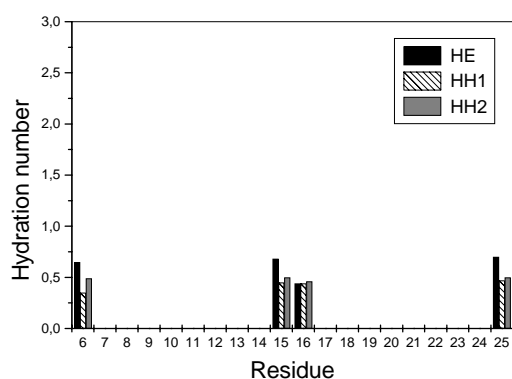


Figure 13. The HB of the arginine side chain HE, HH1 and HH2 hydrogen atoms in the *HNP-3* monomer 1.

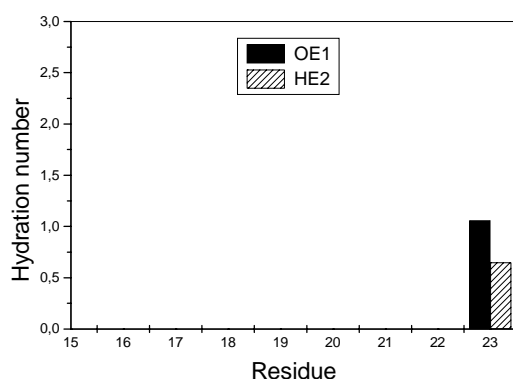


Figure 14. The HB of the OE1 and HE2 glutamine residue atoms in the *HNP-3* monomer 1.

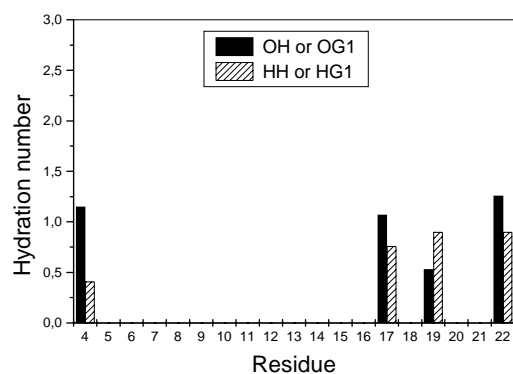


Figure 15. The HB of the OH, HH Tyr and OG1, HG1 Thr residues atoms in the *HNP-3* monomer 1.

The selected  $g_{A,O}(r)$  and  $g_{B,H}(r)$  and corresponding  $P(E)$  profiles (not shown) indicate that the polar side chains of many amino acids present well defined hydration shells. The spatial arrangement of the cationic residues on the peptide surface enhances the action of the hydrophobic forces in the dimer inner region. Hydrophobic interactions are usually thought as resulting from a partial reversal of the solvation process that can be seen as due to the weak exposure or the lack of exposure of apolar groups to the water phase with

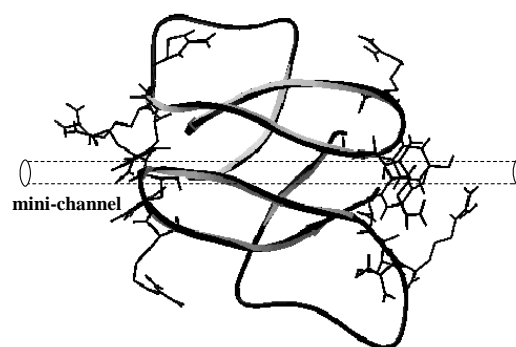


Figure 16. The hydrophobic mini-channel that completely crosses the *HNP-3* dimer.

the consequently unfavorable entropy contribution to the free energy of solvation, is minimized. The *HNP-3* amphiphilic character is probably the key of the affinity of this peptide, and other of the same class, to negatively charged bacterial membranes.

### 4.3 The gramicidin A

The initial structure of the GA used in the simulations was obtained by NMR in dimyristoilphosphatidylcoline bilayer (DMPC)[29], pdb code: 1MAG[60]. The simulated system was constituted by a GA dimer and a mixture of polar (water) and apolar (carbon tetrachloride) solvents. This system intends to mimic internal apolar region of the biological membranes and their polar neighborhoods. The simulation box was divided in three slices containing the apolar phase in the central slice and the polar phases in the two other slices. The central part of the simulation box embodies the GA channel structure excepting the ethanolamine polar groups ( $-NH-CH_2CH_2OH$ ) that must remain in contact with the aqueous phase.

The Fig. 17 represents the last configuration of the system obtained by simulation that presents a  $rmsd$  from the respective experimental starting structure of  $0.16nm$  characterizing the structural conservation of the peptide. The GA dimer forms a structure that passes through the apolar phase connecting the two aqueous phases. This structure is a channel since a row of water molecules occupies its center, Fig. 18. These water molecules are aligned in such a way that they are connected by HB. The water molecule trajectories indicate that their residence time inside the channel is about  $0.26ns$  signaling a constant exchange of water molecules between the internal part of the channel and the water phase. The HB formed by the water molecules, that are inside the channel, with the oxygen atoms of the CO and NH atoms present  $E_{HB}$ , between  $-7.0$  and  $-11.0$  kcal/mol. The  $E_{HB}$  of the oxygen and hydrogen atoms of the hydroxyl groups present in the ethanolamine group are found around  $-9.7$  kcal/mol and  $-15.7$  kcal/mol, respectively. These intermolecular HB are consequently an important stabilizing factor of the channel structure in its active form.

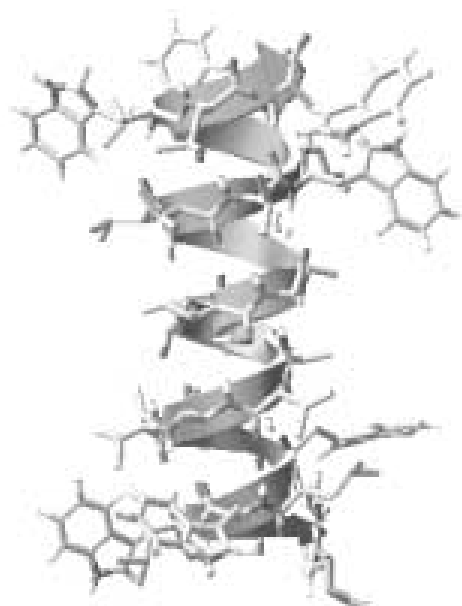


Figure 17. The last structure of the *gramicidin A* obtained by simulation.

The stability of the structure generated during the simulation can be evaluated by monitoring of the intramolecular HB involving the oxygen and hydrogen atoms of the polar groups in all residues. It can be concluded that the GA channel structure is maintained in its active form by the stabilizing effects due to the intermolecular HB with water molecules and by direct interaction between the monomers. The datas relative to the intermonomer HB are presented in Table 1 where it can be seen that the intermonomer HB were detected during more than 70% of the time indicating that the dimer structure exhibits a high stability. These HB stabilize both the internal region of the channel as so as its contacts with the aqueous phase at both ends. The maintenance of the intramolecular HB network is enough to characterize the stability of the channel structure in the active form. The structural stability is a consequence, among other factors, of the intermolecular interactions in which the water molecules carry out the important function of stabilizing the peptide main chain.

Table 1. The HB between the two monomers *gramicidin A* in the GA channel.

CO	NH	$x_{max}$ (Å)	Fr
Val(1)	Ala(20)	1.95	0.96
Ala(5)	Val(16)	1.95	0.75
Ala(3)	Ala(18)	1.95	0.98
Ala(20)	Val(1)	1.95	0.86
Val(16)	Ala(5)	1.95	0.95
Ala(18)	Ala(3)	1.95	0.99



Figure 18. The water molecules that occupy the center of the channel in the last structure of the *gramicidin A* obtained by simulation.

#### 4.4 The myotoxin II

The trajectory of *BaspMT-II* reveals that its structure remains highly stable after 0.5 ns because the *rmsd* oscillates near 3.0nm, Fig. 19, showing also that the secondary structures of the helix-I, II and III are well defined and stable. This stabilization occurs by means of intra and intermolecular HBs. The intramolecular HBs that were found are 112 HBs between H and O backbone atoms, 111 HBs between backbone and side chains atoms and 45 HBs between atoms of side chains. Intermolecular HBs with the backbone atoms were observed with the larger frequency in  $\beta$ -wing and C-terminal regions. The helices are stabilized mainly by intramolecular HB because the number of intramolecular HB is larger than the number of intermolecular HB, Fig. 20.

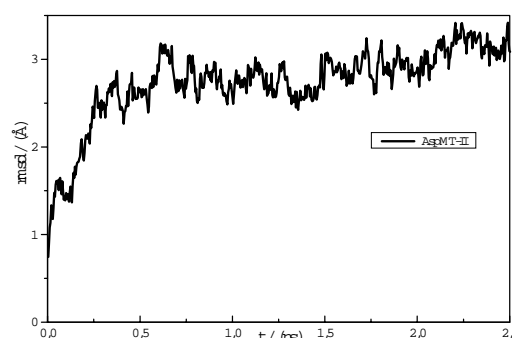


Figure 19. The *rmsd* of the *BaspMT-II*. The reference structure is the initial structure of the *BaspMT-II* used in the simulations.

The stabilization of the helices is better observed from the improvement of the helix-I and II secondary structures according to the Ramachandran dihedral angles criteria[4]. Another point of improvement of the secondary structure is



a new short helix formed by four residues (115-118) with charged side chains located in the C-terminal region. The  $\beta$ -wing and C-terminal loop are the most hydrated regions on account of the presence of charged side chains exposed to the solvent, Fig. 20. The Lys49 charged group  $NH_3^+$  is directed to the calcium binding loop with the three hydrogen atoms hydrogen bonded with three water molecules. In the final structure of the simulation, one of these water molecules is hydrogen bonded with the ND1 atom of the His48 residue. An HB network is observed around these water molecules. These results are supported by the  $n_{HB}$  of the hydrogen atoms of the  $NH_3^+$  group in Lys49 and of the ND1 atom of His48. The water molecule that is hydrogen bonded with the ND1 atom of the His48 residue is an hydration water molecule (W-His48).

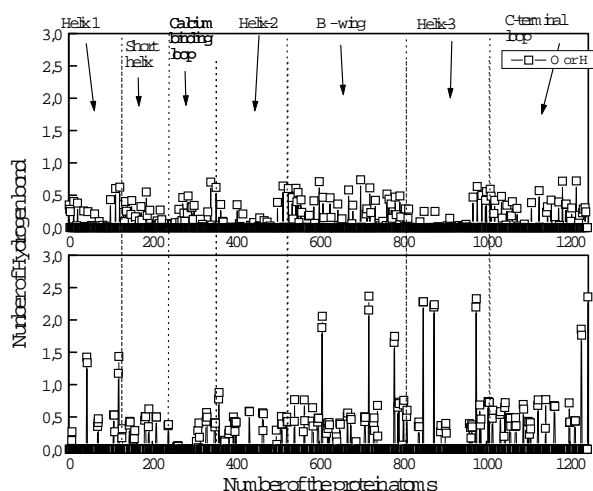


Figure 20. The intermolecular HB between oxygen (O) and hydrogen (H) atoms of *BaspMT-II* of protein and water molecules. The plot shows HB in each of the secondary structure of the *BaspMT-II* monomer. The data of the backbone atoms are plotted in the first part of the figure and the side chain data in the second part.

The HB network, located into hydrophobic channel, has an essential role in the stabilization of the local structure. The water molecule, W-His48, detected in the catalytic site, is the molecule shown into the circle in the Fig. 5. The localization of W-His48 is in agreement with the experimental results[38] and with the mechanism proposed to explain the catalytic activity[35, 38]. Other water molecules can also be identified in the channel forming a HB network with W-His48.

The simulation showed that the W-His48 is present in the Lys49 PLA2s, as is Asp49 PLA2s. Consequently, the presence of W-His48 cannot guarantee the ability to catalyze. The lack of the catalytic activity in the Lys49 PLA2s is explained by the existence of a  $NH_3^+$  group, in place of the co-factor  $Ca^{2+}$  in Asp49 PLA2s, coordinated with calcium binding loop carbonyl oxygen atoms. The  $NH_3^+$  group of

the Lys49 PLA2s residue plays a role in the stability of the local structural but cannot stabilize the oxyanion transition state in the same way as the co-factor  $Ca^{2+}$  can do it. The access of the  $Ca^{2+}$  co-factor is completely hindered because the  $NH_3^+$  group of the Lys122 residue, that is located on the other side of the loop, interacts strongly with oxygen carbonyl atoms of the loop forming intramolecular HBs.

The structure of the *myotoxin II* is stabilized in great part by intramolecular HBs. Parts of the secondary structure are stabilized by intermolecular HBs excepting the helix-2 and helix-3. The local structure of the hydrophobic channel is stabilized by a HB network. The maintenance of the W-His48 plays an important role in the structure and function of both PLA2s, Lys49 PLA2 and Asp49 PLA2. The  $Ca^{2+}$  co-factor is essential to realize the catalyze, but it is not essential to stabilize the structure of the Lys49 PLA2s.

## 4.5 An amphipathy scale

In order to determine a hydrophobicity scale, molecular simulations of small peptides in spc/e water were performed with small peptides like *aa* (1aa), *Gly-aa-Gly* (3aa) and *Gly-Gly-aa-Gly-Gly* (5aa), where *aa* corresponds to the amino acids commonly found in the protein primary structures. The amino acid glycine was used as reference because its side chain is constituted by only one hydrogen atom that is unable to interact strongly with other atoms. The mean configurational energies peptide-water molecules were focused for all the peptides. They can be splitted in two parts: the backbone average configurational energy,  $E_{bb}(iaa)$  with  $i \equiv 1, 3$  or 5, and side chains average configurational energy,  $E_{sc}(iaa)$ .

The values of  $E_{bb}(1aa)$ ,  $E_{bb}(3aa)$  and  $E_{bb}(5aa)$  listed in Table 2 indicate that the backbone configurational energy are somewhat constant since it does not depend on the nature of the central *aa* residue. The  $E_{bb}(1aa)$  average is equal to  $(-113,2 \pm 6,8)$  kcal/mol,  $(-224,8 \pm 6,1)$  kcal/mol for  $E_{bb}(3aa)$  and  $(-231,9 \pm 14,3)$  kcal/mol in the  $E_{bb}(5aa)$  case. Of course, deviations of  $E_{bb}(iaa)$  from the average are observed being the larger deviation observed with the proline amino acid because its side chain is a cycle and the hybridization of the N atom of its main chain is  $sp^2$  not  $sp^3$  as for the other amino acids since the *N*-terminal of proline is an imino not an amino group so that proline would be more correctly classified as an imino acid. In this way, the lack of one hydrogen atom bonded to the nitrogen atom in the proline amino acid results in less attractive  $E_{bb}(iPr o)$ .

From these results, it is obvious that the chemical characteristics of side chain of the residues have only a little influence on the intramolecular interaction energies of the backbone atoms. Until the present, nothing can be said on the side chains configurational energies that, nevertheless, must be fully responsible for the hydrophobicity of amino acids.

Table 2. The backbone configurational energies of 1aa, 3aa and 5aa.

Amino acid	$E_{bb}(1aa)(\text{kcal.mol}^{-1})$	$E_{bb}(3aa)(\text{kcal.mol}^{-1})$	$E_{bb}(5aa)(\text{kcal.mol}^{-1})$
Ala	-119.02	-227.58	-216.69
Arg	-115.17	-226.07	-215.66
Asn	-115.18	-230.49	-221.62
Asp	-102.51	-216.38	-228.68
Cys	-115.91	-222.30	-222.90
Gln	-112.70	-230.20	-240.02
Glu	-108.05	-217.43	-219.01
Gly	-118.91	-229.32	-240.96
His	-112.50	-226.82	-245.44
Ile	-113.54	-221.58	-254.78
Leu	-115.69	-228.73	-249.10
Lys	-122.71	-229.61	-245.68
Met	-112.19	-231.04	-245.68
Phe	-114.00	-222.51	-228.53
Pro	-90.89	-207.99	-204.59
Ser	-114.05	-229.03	-216.08
Thr	-112.75	-217.42	-239.62
Trp	-117.14	-228.67	-249.32
Tyr	-118.54	-223.77	-222.51
Val	-112.30	-229.51	-229.99

Table 3. The side chain configurational energies of 1aa, 3aa and 5aa.

Amino acid	$E_{sc}(1aa)(\text{kcal.mol}^{-1})$	$E_{sc}(3aa)(\text{kcal.mol}^{-1})$	$E_{sc}(5aa)(\text{kcal.mol}^{-1})$
Ala	-1.81	-1.70	-1.59
Arg	-56.50	-63.21	-57.37
Asn	-16.26	-18.67	-18.40
Asp	-90.59	-110.89	-108.19
Cys	-4.48	-4.58	-4.34
Gln	-19.14	-19.45	-18.60
Glu	-112.42	-115.84	-115.94
Gly	0	0	0
His	-20.33	-24.01	-23.55
Ile	-5.91	-5.54	-5.41
Leu	-5.96	-5.84	-5.62
Lys	-76.74	-79.50	-81.48
Met	-6.90	-6.80	-6.67
Phe	-11.74	-11.65	-11.31
Pro	-2.99	-3.95	-3.92
Ser	-9.06	-13.58	-13.68
Thr	-9.68	-13.79	-14.32
Trp	-17.22	-17.26	-17.42
Tyr	-24.19	-24.51	-24.57
Val	-4.62	-4.38	-4.32

Table 4. The side chain average configurational energies.

Amino acid	$\langle E_{sc} \rangle$ (kcal.mol <sup>-1</sup> )	Amino acid	$\langle E_{sc} \rangle$ (kcal.mol <sup>-1</sup> )
Ala	-1.7±0.1	Leu	-5.8±0.2
Arg	-59.1±3.6	Lys	-79.3±2.4
Asn	-17.8±1.3	Met	-6.8±0.1
Asp	-103.2±11.1	Phe	-11.6±0.2
Cys	-4.4±0.1	Pro	-3.6±0.5
Gln	-19.1±0.4	Ser	-12.1±2.6
Glu	-114.7±2.1	Thr	-12.6±2.5
Gly	0	Trp	-17.3±0.1
His	-22.6±2.1	Tyr	-24.4±0.2
Ile	-5.6±0.3	Val	-4.4±0.2

The values of  $E_{sc}(1aa)$ ,  $E_{sc}(3aa)$  and  $E_{sc}(5aa)$  are listed in Table 3. They are quite invariable in function of the number of Gly residues in the peptides *1aa*, *3aa* and *5aa*. The model *1aa*, *3aa* and *5aa* peptides are small so that they are unable to protect their hydrophobic side chains from direct interactions with the solvent. Table 4 lists the average values,  $\langle E_{sc} \rangle$ , of  $E_{sc}(1aa)$ ,  $E_{sc}(3aa)$  and  $E_{sc}(5aa)$ . Consistently with our model,  $\langle E_{sc} \rangle = 0$  for the glycine amino acid where the side chain is a hydrogen atom not explicitly identified in the Gromos96 force field[56]. From a single inspection of Table 4 it can be seen that the amino acids can be grouped in 3 sets: one where the amino acids with energies  $\langle E_{sc} \rangle$  between 0 and -7 kcal/mol are put together constituting the group of the apolar amino acids. The second group aggregates the polar amino acid with

$\langle E_{sc} \rangle$  found in the range -12 to -25 kcal/mol. The last group is the group of charged polar amino acid: their  $\langle E_{sc} \rangle$  energies are in the -60 to -115 kcal/mol range. It is interesting to note that the Gromos96 force field[56] is consistent with the fact that the phenylalanine and tryptophan amino acids are weakly polar so that a charge less than  $0.2e$  is attributed to the aromatic hydrogen atoms. Moreover, the three-dimensional structure of the proteins depends not solely on the hydrophobicity of the side chains but depends also on the steric effects due to the different extensions of the space hindered by the different secondary structures in the interior of proteins[61, 62].

As a final classification, the results listed in table 4 can provide the yearned hydrophobicity sequence of the 20 amino acids:

---

largest hydrophobicity  $\rightarrow$  Gly  $\rightarrow$  Ala  $\rightarrow$  Pro  $\rightarrow$  (Cys, Val)  $\rightarrow$  (Ile, Leu)  $\rightarrow$  Met  $\rightarrow$  Phe  $\rightarrow$  (Ser, Thr)  
 $\rightarrow$  Trp  $\rightarrow$  Asn  $\rightarrow$  Gln  $\rightarrow$  His  $\rightarrow$  Tyr  $\rightarrow$  Arg  $\rightarrow$  Lys  $\rightarrow$  (Asp, Glu)  $\rightarrow$  largest hydrophilicity

---

The  $\langle E_{sc} \rangle$  of the amino acids put together are very close. The  $\langle E_{sc} \rangle$  side chain energies configurational present some advantages when compared with other amphipathy scales:

1. they consider the amino acids when they are bonded with other residues forming peptide chains;
2. they analyse the complete interaction between the solvent and the amino acids (backbone and side chain);
3. they are able to reproduce exactly the environment amino acid-water;
4. they do not depend on unclear differences of standard free energies;
5. they do not depend on hypothesis frequently included in many amphipathy scale determination.

Nevertheless, the quality of the present results has to be confirmed by analyzing the interaction energies between the

solvent and the side chains of proteins in aqueous phases and comparing these results with present  $\langle E_{sc} \rangle$  scale. For this analysis, the next proteins used were: *basic fibroblast growth factor*[63], *defensin*[26], *gramicidin*[64] and two *conotoxins*[65, 66]. The mean side chain-solvent configurational energies obtained with these proteins will be called  $\langle E_{proteins} \rangle$ . The correlation between  $\langle E_{proteins} \rangle$  and  $\langle E_{sc} \rangle$  is presented in Fig. 21 where the correlation coefficient is equal to 0.97. The absolute values of  $\langle E_{sc} \rangle$  are always larger than the absolute values of  $\langle E_{proteins} \rangle$  because the access of the solvent to the side chains is always more hindered in the real structures than in the *1aa*, *3aa* and *5aa* model peptides. Consequently, the solvent molecules are always localized at distances such that the configurational intermolecular energies are always less attractive in the protein case. The intermolecular configurational energies for the side chain of each residue of the *basic fibroblast growth factor* ( $E_{1BFF}$ ) are plotted in Fig. 22 as a function of the position of the residue. It can be noted that some re-

gions present less attractive energies, they are specifically the regions 36-40, 61-65, 69-73, 108-114, 119-126 and 144-152. It is interesting to observe that these regions are hydrophobic regions localized in the interior of *basic fibroblast growth factor*. The regions shown in Fig. 23 are hydrophilic regions situated on the external part of the *basic fibroblast growth factor*. It is, consequently, clear that parts of secondary structures can be predicted from the  $\langle E_{sc} \rangle$  scale.

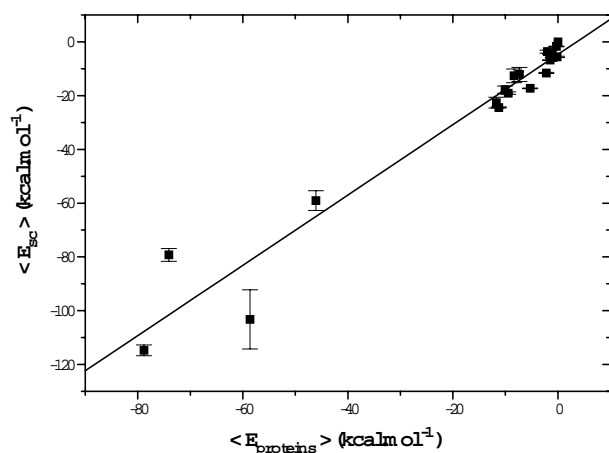


Figure 21. The plot of  $\langle E_{proteins} \rangle$  vs.  $\langle E_{sc} \rangle$ . The correlation coefficient is equal to 0.97.

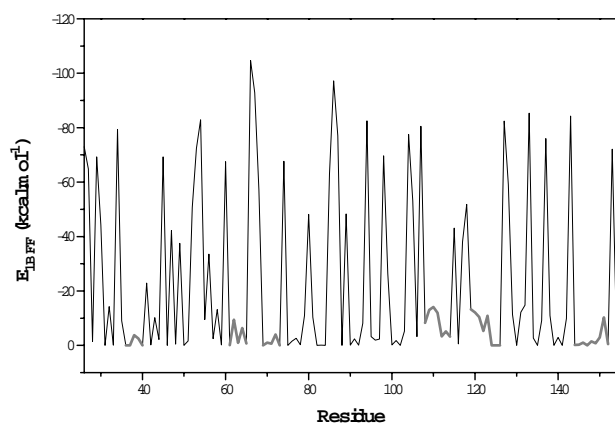


Figure 22. The sequence of the  $\langle E_{sc} \rangle$  in the *fibroblast growth factor* primary structure. The gray regions are the hydrophobic regions.

## 5 Conclusion

In this paper, important features of the role of water in the protein activity were presented and discussed. The fundamental function of the intermolecular HB in the tertiary and quaternary protein structures was described focusing the maintenance of local structures as so as the insertion of the peptides into the larger, and very stable, HB network formed by the interactions between the solvent molecules. The examples proposed to enlight the water importance are clear

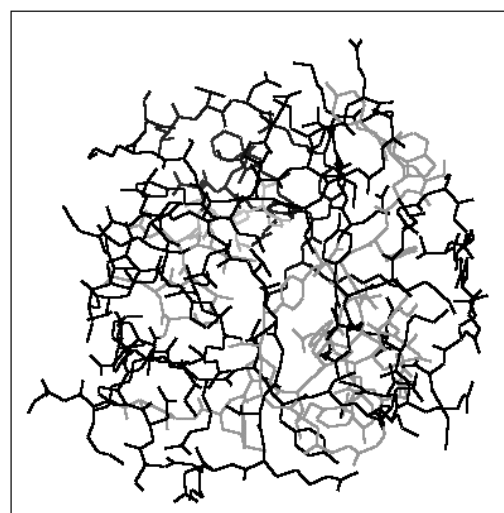


Figure 23. The position of hydrophobic regions detected in Fig. 22 in the *fibroblast growth factor* crystal structure.

in the aspects that HB are of main importance considering that:

1. the *basic fibroblast growth factor* is unable to define its tertiary structure without the solvent contribution since only  $\beta$ -strands, and no disulfide bridge, are present in its structure;
2. the importance of the water molecules is enhanced in considering the dimer *defensin* where the hydrophobic channel is stabilized by water molecules at both ends. The immediate consequence is that highly reactive polar and charged side chain groups are firmly exposed to external interactions;
3. the same stabilizing influence of the water molecules are observed in the case of the *gramicidin* channel since both interior region and mouths suffer the influence of the water that can pass through the channel as it must occur as well when *gramicidin* forms a channel in bacterial membranes resulting in the loss of protoplasmic materials and in the death of the bacteria;
4. the double function of a water molecule in the maintenance of the catalytic center structure and in the hydrolysis of the sn-2 acyl bonds of sn-3 phospholipids is outstanding in the *phospholipase* study;
5. finally a method based on objective simulation data has permitted to predict secondary peptide structures that are fundamental features in the protein activities.

All these examples on the role of the water molecules permit to confirm that the living systems are like they are because water is an outstanding and abundant molecule present everywhere in living matter. This is the reason why the water molecule can be baptized as the most important component in the living matter.

## Acknowledgments

This work was supported in part by the Conselho Nacional de Desenvolvimento Científico e Tecnológico and by the Fundação de Amparo à Pesquisa do Estado de São Paulo.

## References

- [1] C. Branden, J. Tooze, *Introduction to Protein Structure*, Garland Publishing, New York (1999).
- [2] A. L. Lehninger, D. L. Nelson, and M. M. Cox, *Princípios de Bioquímica*, Sarvier, São Paulo (1995).
- [3] L. Stryer, *Biochemistry*, W. H. Freeman and Company, New York (1995).
- [4] D. Voet, J. Voet, *Biochemistry*, John Wiley and Sons, New York (1995).
- [5] J. Darnell, H. Lodish, and D. Baltimore, *Molecular Cell Biology*, Scientific American Books, New York (1986).
- [6] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, *Molecular Biology of The Cell*, Garland Science, New York (2002).
- [7] J. Étienne, *Bioquímica Genética e Biologia Molecular*, Livraria Santos Editora, São Paulo (2003).
- [8] G. A. Jeffrey, W. Saenger, *Hydrogen Bonding in Biological Structures*, Springer-Verlag (1994).
- [9] R. B. Gregory ed., *Protein-solvent interactions*, Marcel Dekker, New-York (1995)
- [10] L. Blum, L. Degrève, *Mol. Phys.* **88**, 585 (1996).
- [11] L. Degrève, L. Blum, *Physica A* **224**, 550 (1996).
- [12] J. Zhang, L. S. Cousens, P. J. Barr, S. R. Sprang, *Proc. Natl. Acad. Sci. USA*, **88**, 3446 (1991).
- [13] G. G. Gallego, G. Conn, V. B. Hatcher, K. A. Thomas, *Biochem. and Biophysical Res. Comm.* **135**, 541 (1986).
- [14] H. Ago, Y. Kiagawa, A. Fujishima, Y. Matsuura, and Y. Katsube, *J. Biochem.* **110**, 360 (1991).
- [15] B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, and J. D. Watson, *Biologia Molecular da Célula*, Artes Médicas, Porto Alegre, 1997.
- [16] F. J. Moy, A. P. Seddon, P. Bohlen, and R. Powers, *Biochemistry*, **35**, 13552 (1996).
- [17] R. E. W. Hancock, R. Lehrer, *Trends Biotechnol.* **16**, 82 (1998).
- [18] R. M. Epanand, H. J. Vogel, *Biochim. Biophys. Acta* **1462**, 11 (1999).
- [19] J. Patterson-Delafield, D. Szklarek, R. J. Martinez, and R. I. Lehrer, *Infect. Immun.* **31**, 723 (1981).
- [20] R. I. Lehrer, A. Barton, K. A. Daher, S. S. L. Harwig, T. Ganz, and M. E. Selsted, *J. Clin. Invest.* **84**, 553 (1989).
- [21] S. H. White, W. C. Wimley, and M. E. Selsted, *Curr. Opin. Struct. Biol.* **5**, 521 (1995).
- [22] Y. Shai, *Trends Biochem. Sci.* **20**, 460 (1995).
- [23] T. Ganz, M. E. Selsted, D. Szklarek, S. S. L. Harwig, K. Daher, D. F. Bainton, and R. I. Lehrer, *J. Clin. Invest.* **76**, 1427 (1985).
- [24] T. Ganz, R. I. Lehrer, *Curr. Opin. Immun.* **6**, 584 (1994).
- [25] T. Ganz, R. I. Lehrer, *Pharmac. Ther.* **66**, 191 (1995).
- [26] C. P. Hill, J. Yee, M.E. Selsted, and D. Eisenberg, *Science*, **251**, 1481 (1991).
- [27] P. Yeagle, *The Structure of Biological Membranes*, CRC Press, (1992).
- [28] R. B. Gennis, *Biomembranes - Molecular Structure and Function*, Springer-Verlag, New York, (1989).
- [29] R. R. Ketchum, W. Hu, T. A. Cross, *Science*, **261** (10), 1457 (1993)
- [30] B. A. Wallace, *Annu. Rev. Biophys. Biophys. Chem.* **19**, 127 (1990).
- [31] W. R. Veatch, E. T. Fossel, and E. R. Blout, *Biochemistry*, **13**, 5249 (1974).
- [32] W. R. Veatch, E. R. Blout, *Biochemistry*, **13**, 5257 (1974).
- [33] B. E. Isbell, C. Rice-Evans, and G. H. Beaven, *FEBS LETTERS*, **25**, 192 (1972).
- [34] E. T. Fossel, W. R. Veatch, Y. A. Ovchinnikov, and E. R. Blout, *Biochemistry*, **13**, 5264 (1974).
- [35] R. K. Arni, R. J. Ward, *Toxicol.* **34**, 827 (1996).
- [36] E. A. Dennis, *J. Biol. Chem.* **269**, 13.057 (1994).
- [37] P. Rosenberg, in: W. T. Shier, D. Mebs (Eds.), *Handbook of Toxicology*, Marcel Dekker, New York, (1990).
- [38] D. L. Scott, S. P. White, Z. Otwinowski, W. Yuan, M. H. Gelb, and P. B. Sigler, *Science* **250**, 1545 (1990).
- [39] C. J. Van den Bergh, A. J. Slotboom, H. M. Verheij, and G. H. De Haas, *J. Cell. Biochem.* **39**, 379 (1989).
- [40] Y. Li, B. Yu, H. Zhu, M. Jain, and M. Tsai, *Biochemistry* **33**, 14714 (1994).
- [41] Y. Nozaki, C. Tanford, *J. Biol. Chem.* **246**, 2211 (1971).
- [42] S. Damodaran, K. B. Song, *J. Biol. Chem.* **261**, 7220 (1986).
- [43] S. H. White, W. C. Wimley, *Annu. Rev. Biophys. Biomol. Struct.* **28**, 319 (1999).
- [44] C. Chothia, *J. Mol. Biol.* **105**, 1 (1976).
- [45] P. Manavalan, P. K. Ponnuswamy, *Nature*, **275**, 673 (1978).
- [46] S. Fraga, *Can. J. Chem.* **60**, 2606 (1982).
- [47] R. M. Sweet, D. Eisenberg, *J. Mol. Biol.* **171**, 479 (1983).
- [48] L. A. Kuhn, J. S. Leigh, *Biochim. Biophys. Acta*, **828**, 351 (1985).
- [49] H. Tanaka, *J. Chem. Phys.* **86**(3), 1512 (1987).
- [50] G. Hummer, S. Garde, A. E. Garcia, M. E. Paulaitis, and L. R. Pratt, *J. Phys. Chem. B* **102**(51), 10469 (1998).
- [51] W. Blokzijl, J. B. F. N. Engberts, *Angew. Chem. Int. Ed. Engl.* **32**, 1545 (1993).
- [52] C. Chothia, *Nature*, **248**, 338 (1974).
- [53] B. J. Alder, T. E. Wainwright, *J. Chem. Phys.* **31**, 459 (1959).
- [54] G. Cicotti, D. Frenkel, and I. R. McDonald, *Simulation of Liquids and Solids*, North-Holland, Amsterdam, (1990).
- [55] M. P. Allen, D. J. Tildesley, *Computer Simulation of Liquids*, Clarendon Press, Oxford, (1987).

- [56] W. F. van Gunsteren, S. R. Billeter, A. A. Eising, P. H. Hünenberger, P. Krüger, A. E. Mark, W. R. P. Scott, and I. G. Tironi, *Biomolecular Simulation: The GROMOS 96 Manual and User Guide*, Biomos, Groningen, (1996).
- [57] I. G. Tironi, R. Sperb., P. E. Smith, and W. F. van Gunsteren, *J. Chem. Phys.* **102**, 5451 (1995).
- [58] F. H. Stillinger, A. Rahman, *J. Chem. Phys.* **60**, 1545 (1974).
- [59] L. B. Silva, L. Degrève, *Mol. Phys.* **100**, 3111 (2002).
- [60] <http://www.rcsb.org/pdb>
- [61] S. Lifson, C. Sander, *Nature*, **282**, 109 (1979).
- [62] J. Janin, C. Chothia, *J. Mol. Biol.* **143**, 95 (1980).
- [63] J. S. Kastrup, E. S. Eriksson, H. Dalboge, and H. Flodgaard, *Acta Crystallogr. D Biol. Crystallogr.* **53**, 160 (1997).
- [64] R. R. Ketchum, K. C. Lee, S. Huo, and T. A. Cross, *J. Biomol. NMR*, **8**, 1 (1996).
- [65] S. Farr-Jones, G. P. Miljanich, L. Nadasdi, J. Ramachandran, and V.J. Basus, *J. Mol. Biol.* **248**, 106 (1995).
- [66] T. Kohno, J. I. Kim, K. Kobayashi, Y. Kodera, T. Maeda, and K. Sato, *Biochemistry*, **34**, 10256 (1995).