

## 2-Dimensional HP Foldings of Dermaseptin-J2

Shaomin Yan, Guang Wu

State Key Laboratory of Non-Food Biomass Enzyme Technology, National Engineering Research Center for Non-Food Biorefinery, Guangxi Key Laboratory of Biorefinery, Guangxi Academy of Sciences, Nanning, China  
Email: hongguanglishibahao@yahoo.com

Received November 2012

### ABSTRACT

Although the hydrophobic-polar (HP) model is a simple model to study protein folding, it is an approximation to the real-life case. Dermaseptin is a subfamily of frog skin active peptide family, which has various antimicrobial activities, and dermaseptin-J2 is a newly found peptide composed of 26 amino acids. In this study, the 2-dimensional HP model was used to analyze the foldings of dermaseptin-J2 and its nine mutants, which were converted to different HP sequences according to the normalized amino acid hydrophobicity index with respect to pH levels and the conversion of glycine as hydrophobic or polar, and each has 847,288,609,443 possible foldings. The results show that the foldings with minimal energy have different native states, which are chiral and can be numerically distinguished and ranked according to the normalized amino acid hydrophobicity index. The nine mutants of dermaseptin-J2 do not affect the minimal energy but affect their native states at pH 7. The results demonstrate that two pH levels and conversion of glycine as hydrophobic or polar affect the native state and minimal energy, suggesting these are two ways to modify dermaseptin-J2.

**Keywords:** Dermaseptin; Folding Configuration; HP Model; Hydrophobicity Index; Minimal Energy; Native State

### 1. Introduction

Protein folding is important to understand its structure-function relationship and folding process. The hydrophobic-polar (HP) model is a very simple model, which was based on the observation that hydrophobic interaction was the driving force for protein folding and the hydrophobicity of amino acids was the main force for developing a native conformation of small globular proteins [1].

The HP model suffers from critiques because its assumption was simple for protein folding, its description was far away from real-life case, its results were difficult to cooperate with the folding obtained from experiments, etc. Therefore, the HP model is not a model without limitations, however, any model is an approximation to the real-life case, and the difference between models relies on the degree of their approximations. If one takes a viewpoint from a single model, then all of the rest models will be problematic.

Actually, any model helps us to understand the world from a different angle. There are several reasons to study the HP model in great details: i) so far very few studies using HP model were applied to real-life case because the HP model needs extremely intensive computations, therefore the studies using HP model is a way to test our computing ability; ii) the extremely intensive computa-

tions in HP model was classified as NP problem [2], being the first problem listed millennium prize, so the study on HP model is approaching to one of numerous unsolved examples of NP problem; iii) the study on HP model would help us develop optimal algorithms [3], which sheds light on solving intensively computational problems in biological branches such as phylogenetics, RNA pseudoknot [4]; iv) the HP model provides different insights into folding process; etc.

The HP model is workable for 2-dimensional (2D) and 3-dimensional (3D) folding. For both cases, each amino acid, either hydrophobic (H) or polar (P), walks along a line in 2D lattice or in 3D cube by taking a self-avoided step, then an H-H connection, which does not come from sequential step, has minus unity energy [5], and the folding with minimal energy would be a protein's native folding.

Current computing power can afford the studies on folding in a 2-dimensional HP model for very short protein. Dermaseptin is a subfamily of frog skin active peptide family, which has various antimicrobial activities. Dermaseptin-J2 was a relatively newly found peptide of 26 amino acids [6].

The number of foldings increases dramatically with the increase in the length of protein sequence. For example, each amino acid would have three directions to go in self-avoided step in 2D lattice, so the number of three

direction steps for  $n$  amino acids would be  $3^{(n-1)}$ . In the case of 26-residue dermaseptin-J2, the possible foldings are 847,288,609,443 in 2D HP model. Practically, it is useful to know all the possible foldings of protein. The aim of this study is to use the 2D HP model to analyze all possible foldings of dermaseptin-J2 with hope to get insight into dermaseptin-J2 antibiotic activity.

## 2. Materials and Methods

### 2.1. Data

The amino acid sequence of dermaseptin-J2 was obtained from UniProt [7], and its accession number was P86636. The normalized amino acid hydrophobicity index was obtained from SigmaAldrich website [8].

### 2.2. HP Model

The HP model classifies amino acids as either hydrophobic or polar, but there is no indication regarding neutral amino acids. So we use the normalized amino acid hydrophobicity index [8] to assign amino acids in dermaseptin-J2 either as hydrophobic or as polar, however, this assignment is still not sufficient because this normalized amino acid hydrophobicity index is based on the fact that glycine as zero, thus we have to choose glycine either as hydrophobic or as polar. This leads to an amino-acid sequence of dermaseptin-J2 to have two HP sequences in terms of assigning glycine as hydrophobic or as polar. Again, the amino acid hydrophobicity is pH dependent [8], which leads us to consider the assignment of amino acids of dermaseptin-J2 at two pH levels. Taken two considerations together, one amino-acid sequence of dermaseptin-J2 has four HP sequences to be operated in HP model.

There are nine uncertainties in amino-acid sequence of dermaseptin-J2 [9], *i.e.* L2I, K4Q, L7I, I10L, K12Q, L13I, L19I, K23Q, and L25I. So, we have totally 40 HP sequences for HP model (Table 1), and each one theoretically has 847,288,609,443 foldings.

## 3. Results and Discussion

Currently we have no ability to compute every folding in proteins longer than 30 amino acids. In this study, the Lenovo ThinkPat laptop with due CPU of 2 GHz computed 200,000 to 250,000 foldings per second, for a 26-amino-acid dermaseptin-J2, the computing time was between 39 and 49 days.

With the normalized amino acid hydrophobicity index [8], we can find how neutral amino acids affect the HP sequences (Table 1), where nine mutants have no effects on HP sequences, so their HP foldings are identical implying harmless mutations.

Of 847,288,609,443 possible foldings for each HP se-

**Table 1. Amino acids of dermaseptin-J2 in HP sequences.**

Dermaseptin-J2	Classification	Sequence
Original	Amino acid	glwknmlsgigklageaalgavktlv
	G=H at pH2	hhhpphhphhhphhhhhhhhhhhhhhhhhhh
	G=P at pH2	phhpphhpphphhhphhhhhhhhhhhhhhh
	G=H at pH 7	hhhpphhphhhphhhhhhhhhhhhhhhhhhh
L2I mutant	G = P at pH 7	phhpphhpphphhhpphphhhhhhhhhhhhh
	Amino acid	giwknmlsgigklageaalgavktlv
	G=H at pH2	hhhpphhphhhphhhhhhhhhhhhhhhhhhh
	G=P at pH2	phhpphhpphphhhphhhhhhhhhhhhhhh
K4Q mutant	G=H at pH 7	hhhpphhphhhphhhhhhhhhhhhhhhhhhh
	G=P at pH 7	phhpphhpphphhhpphphhhhhhhhhhhhh
	Amino acid	glwqnmlsgigklageaalgavktlv
	G=H at pH2	hhhpphhphhhphhhhhhhhhhhhhhhhhhh
L7I mutant	G=P at pH2	phhpphhpphphhhphhhhhhhhhhhhhhh
	G=H at pH 7	hhhpphhphhhphhhhhhhhhhhhhhhhhhh
	G=P at pH 7	phhpphhpphphhhpphphhhhhhhhhhhhh
	Amino acid	glwknmlsgigklageaalgavktlv
I10L mutant	G=H at pH2	hhhpphhphhhphhhhhhhhhhhhhhhhhhh
	G = P at pH2	phhpphhpphphhhphhhhhhhhhhhhhhh
	G=H at pH 7	hhhpphhphhhphhhhhhhhhhhhhhhhhhh
	G = P at pH 7	phhpphhpphphhhpphphhhhhhhhhhhhh
K12Q mutant	Amino acid	glwknmlsgigklageaalgavktlv
	G=H at pH 2	hhhpphhphhhphhhhhhhhhhhhhhhhhhh
	G=P at pH 2	phhpphhpphphhhphhhhhhhhhhhhhhh
	G=H at pH 7	hhhpphhphhhphhhhhhhhhhhhhhhhhhh
L13I mutant	G=P at pH 7	phhpphhpphphhhpphphhhhhhhhhhhhh
	Amino acid	glwknmlsgigklageaalgavktlv
	G=H at pH 2	hhhpphhphhhphhhhhhhhhhhhhhhhhhh
	G=P at pH 2	phhpphhpphphhhphhhhhhhhhhhhhhh
L19I mutant	G=H at pH 7	hhhpphhphhhphhhhhhhhhhhhhhhhhhh
	G=P at pH 7	phhpphhpphphhhpphphhhhhhhhhhhhh
	Amino acid	glwknmlsgigklageaalgavktlv
	G=H at pH 2	hhhpphhphhhphhhhhhhhhhhhhhhhhhh
K23Q mutant	G=P at pH 2	phhpphhpphphhhphhhhhhhhhhhhhhh
	G=H at pH 7	hhhpphhphhhphhhhhhhhhhhhhhhhhhh
	G=P at pH 7	phhpphhpphphhhpphphhhhhhhhhhhhh
	Amino acid	glwknmlsgigklageaalgavqtlv
L25I mutant	G=H at pH 2	hhhpphhphhhphhhhhhhhhhhhhhhhhhh
	G=P at pH 2	phhpphhpphphhhphhhhhhhhhhhhhhh
	G=H at pH 7	hhhpphhphhhphhhhhhhhhhhhhhhhhhh
	G=P at pH 7	Phhpphhpphphhhpphphhhhhhhhhhhhh

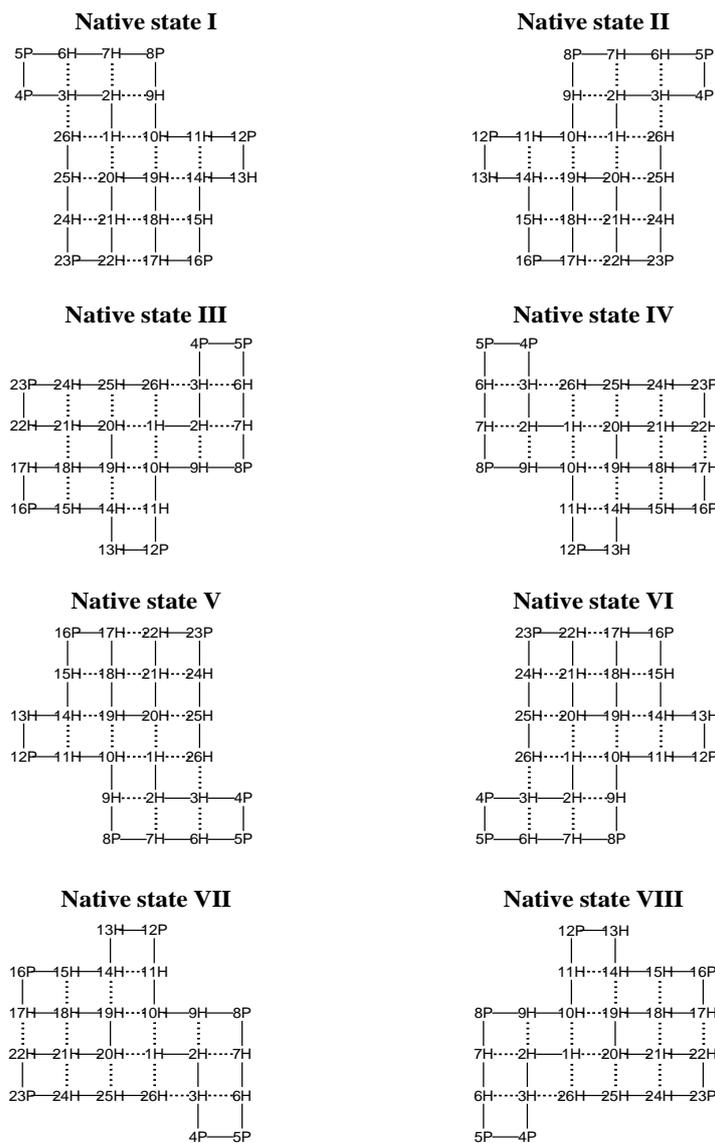
quence listed in **Table 1**, it was important to know how many native states the dermaseptin-J2 had (**Table 2**), where dermaseptin-J2 has more than one native state, for example, 12 native states are found at pH 2 with glycine assigned as polar, and each native state has the same amount of minimal energy,  $-13$ . The fact that there is more than one native state suggests the flexibility of folding mechanisms in dermaseptin-J2.

**Figure 1** shows only 8 native states of dermaseptin-J2 at pH 7 with glycine assigned as polar, in order to have a full picture on the folding process in 2D lattice. In any folding, it begins from position 1 to position 26, which can be viewed as a pathway to form a folding. An interesting point is that the native state is chirally symmetric between the left-hand side and the right-hand side,

namely, the pathways to construct the same folding are chiral because they cannot be superimposed in mirror image. The chiral symmetry in **Figure 1** suggests that all foldings of proteins have possibly chiral symmetry in 2D HP lattice. Therefore a protein can find its native state

**Table 2. Number of native states of foldings with minimal energy of dermaseptin-J2 according to different HP conversions.**

Conversion	Native States	Minimal Energy
G=H at pH 2	30	$-16$
G=P at pH 2	12	$-13$
G=H at pH 7	54	$-15$
G=P at pH 7	72	$-12$



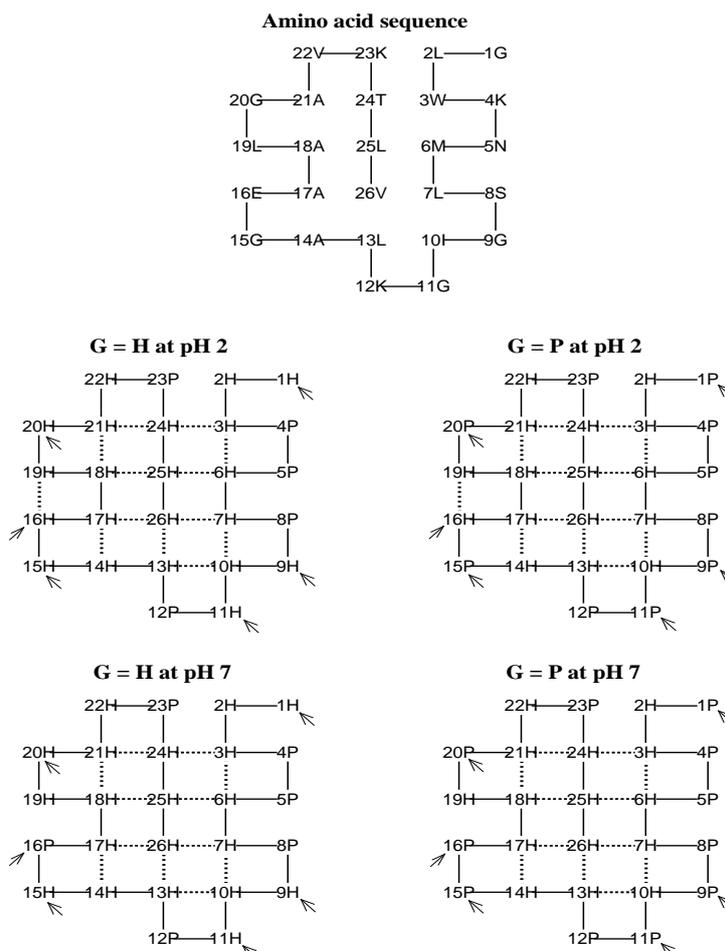
**Figure 1. 8 native states of folding under HP conversion of G=H at pH 7. Dotted lines are non-sequential H-H connection, which is considered as a unit of negative energy  $-1$ , and their sum is the minimal energy  $-15$ .**

through different pathways, which minimize the time spending on searching for the native state.

On the other hand, we notice that the native states and the minimal energy in **Table 2** are different with respect to pH levels and the conversion of glycine as hydrophobic or polar. **Figure 2** shows the amino acid sequence of dermaseptin-J2 and four foldings at pH 2 and pH 7 with glycine assigned as hydrophobic as well as polar. But these four foldings have the same pathway. Here, we need to see the non-sequential H-H connections due to two pH levels and two assignments of glycine. At pH 2 with glycine assigned as hydrophobic (left-hand configuration in middle panel in **Figure 2**), there are 13 non-sequential H-H connections: 3H - 6H, 3H - 24H, 6H - 25H, 7H - 10H, 7H - 26H, 10H - 13H, 13H - 26H, 14H - 17H, 16H - 19H, 17H - 26H, 18H - 21H, 18H - 25H, and 21H - 24H, whose minimal energy is -13 and larger than that of the native states in **Table 2**. At pH 2 with glycine assigned as polar (right-hand configuration in middle panel in **Figure 2**), there are 13 non-sequential H-H connections, whose minimal energy is -13 and equal to that

of the native states in **Table 2**. At pH 7 with glycine assigned as hydrophobic (left-hand configuration in lower panel in **Figure 2**), there are 12 non-sequential H-H connections, whose minimal energy is -12 and larger than that of the native states in **Table 2**. At pH 7 with glycine assigned as polar (right-hand configuration in lower panel in **Figure 2**), there are also 12 non-sequential H-H connections, whose minimal energy is -12 and equal to that of the native states in **Table 2**. Glycines in dermaseptin-J2 marked by arrows are located at edge of HP folding, so they generally do not construct H-H connections with internal Hs, we would expect more dramatic difference if glycines are located in internal part of protein.

Now let us look at **Table 2** again, where a native state has different numbers of folding, such as 40 folding confirmations at pH 2 with glycine assigned as hydrophobic. An intriguing question raised here is whether we can numerically distinguish and rank those folding confirmations? This question comes from such a consideration that a non-sequential H-H connection only gives a unit of



**Figure 2.** Dermaseptin-J2 sequence and four HP foldings at pH 2 and pH 7 with glycines assigned as hydrophobic as well as polar.

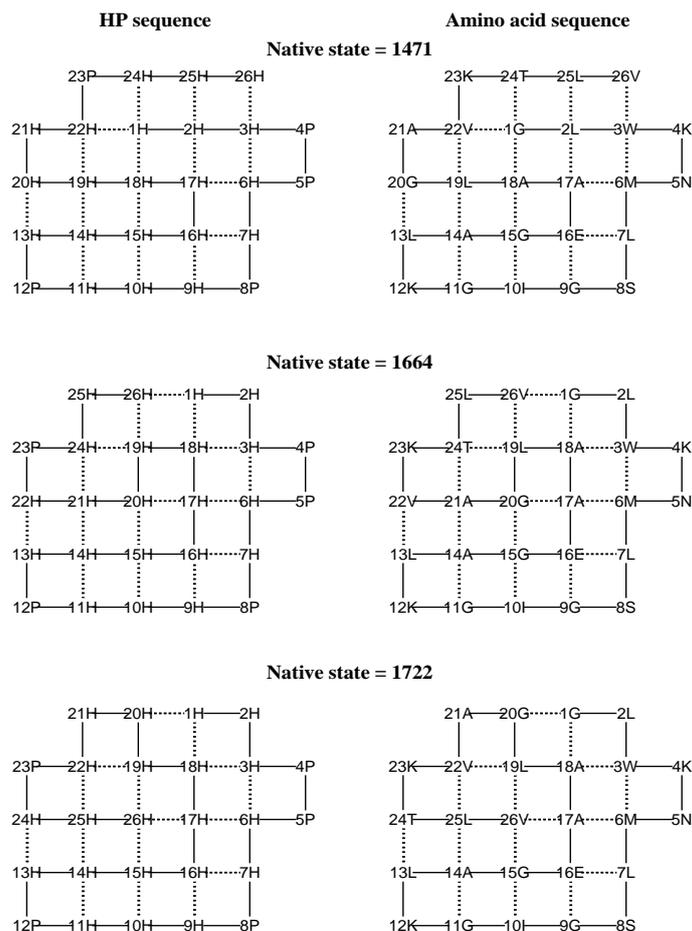
minimal energy, however, an H-H connection is composed of amino acids with different hydrophobicity, and therefore it would be important to further quantify non-sequential H-H connections with normalized amino acid hydrophobicity index [9]. **Figure 3** shows the use of the normalized amino acid hydrophobicity index to numerically distinguish and rank 40 foldings of native state with the same minimal energy  $-16$ , we thus classify those 40 foldings of native state with the same minimal energy  $-16$  into three groups with different foldings.

Furthermore, **Table 3** lists the native state of folding with the minimal energy when the glycine is assigned as hydrophobic or polar at pH 2 or pH 7, including dermaseptin-J2 and its nine mutations. Clearly, those mutants do not have the effects on the folding with minimal energy but have effects on their native states to different degrees. An important observation is the fewer the Hs in an HP sequence, the more the foldings with minimal energy. For an antibiotic, if the number of foldings could be related to the range of antibacterial spectrum, then the state with minimal energy could be related to the speci-

ficity. On the other hand, the lower specificity would imply less potency against targets.

Experiments revealed that some dermaseptin has an inherent propensity to an extended conformation in aqueous solution and self-assembles into amyloid fibrils in a reversible pH-controlled fashion [10]. Implication of this study is one should increase the number of H in order to decrease the minimal energy in a native state. Practically, we can use a hydrophobic amino acid to replace a neutral or hydrophilic amino acid to get a native state with lower minimal energy if this native state concerns chemical reactions.

Finally, the results suggest the possible ways to modify dermaseptin-J2 that we can either modify dermaseptin-J2 via replacing polar amino acids with hydrophobic amino acids or modify dermaseptin-J2 via replacing amino acids according to the normalized amino acid hydrophobicity index. This is meaningful because the antibacterial activity of dermaseptin depends markedly on a threshold number of hydrophobic residues to be present on both extremities of the helix [11].



**Figure 3. Native states under G=H at pH 2 with normalized amino acid hydrophobicity index. All has a minimal energy  $-16$ , which is the sum of dotted lines and different sum of hydrophobicity index of H-H connections. Left-hand site: HP sequence; Right-hand site: Dermaseptin-J2 sequence.**

**Table 3. Native states classified according to normalized amino acid hydrophobicity index.**

Group	State Number	Sum of hydrophobicity index of H-H connections (native states)									
		Original	L2I	K4Q	L7I	I10L	K12Q	L13I	L19I	K23Q	L25I
G=H at pH 2	12	1471	1471	1471	1471	1471	1471	1471	1471	1471	1471
	6	1664	1664	1664	1664	1664	1664	1664	1664	1664	1664
	12	1722	1722	1722	1722	1722	1722	1722	1722	1722	1722
G=P at pH 2	12	1816	1816	1816	1816	1816	1816	1816	1816	1816	1816
G=H at pH 7	6	1393	1397	1393	1395	1391	1391	1393	1397	1393	1393
	6	1548	1552	1548	1550	1544	1544	1550	1552	1548	1548
	6	1553	1557	1553	1557	1551	1551	1553	1557	1553	1553
	6	1566	1568	1566	1568	1564	1564	1566	1570	1566	1570
	6	1576	1580	1576	1578	1572	1572	1576	1580	1576	1578
	6	1708	1712	1708	1712	1704	1704	1710	1712	1708	1708
	6	1721	1723	1721	1723	1717	1717	1723	1725	1721	1725
	6	1726	1728	1726	1730	1724	1724	1726	1730	1726	1730
	6	1881	1883	1881	1885	1877	1877	1883	1885	1881	1885
	6	1881	1883	1881	1885	1877	1877	1883	1885	1881	1885
G=P at pH 7	12	1663	1663	1663	1667	1659	1659	1667	1663	1663	1667
	6	1686	1688	1686	1690	1682	1682	1690	1686	1686	1690
	12	1723	1725	1723	1727	1719	1719	1727	1725	1723	1727
	12	1803	1805	1803	1807	1799	1799	1807	1807	1803	1805
	12	1810	1810	1810	1814	1806	1806	1814	1814	1810	1814
	6	1833	1835	1833	1837	1829	1829	1837	1837	1833	1837
	12	1922	1924	1922	1926	1918	1918	1926	1926	1922	1926
	12	1922	1924	1922	1926	1918	1918	1926	1926	1922	1926

The dermaseptin super-family encompasses a wide variety of structural motifs, and combined approaches have been used to elucidate their antimicrobial effects based on biophysical and cellular biology methods [12]. This study helps to understand dermaseptin folding in terms of HP model.

#### 4. Acknowledgements

This study was partly supported by Guangxi Science Foundation (12237022 and 2013GXNSFDA019007).

#### REFERENCES

- [1] K. A. Dill, "Theory for the Folding and Stability of Globular Proteins," *Biochemistry*, Vol. 24, 1985, pp. 1501-1509. <http://dx.doi.org/10.1021/bi00327a032>
- [2] B. Berger and T. Leight, "Protein Folding in the Hydrophobic-Hydrophilic (HP) Model Is NP-Complete," *Journal of Computational Biology*, Vol. 5, 1998, pp. 27-40. <http://dx.doi.org/10.1089/cmb.1998.5.27>
- [3] T. C. Beutler and K. A. Dill, "A Fast Conformational Search Strategy for Finding Low Energy Structures of Model Proteins," *Protein Science*, Vol. 5, 1996, pp. 2037-2043. <http://dx.doi.org/10.1002/pro.5560051010>
- [4] R. B. Lyngsø and C. N. Pedersen, "RNA Pseudoknot Prediction in Energy-Based Models," *Journal of Computational Biology*, Vol. 7, 2000, pp. 409-427. <http://dx.doi.org/10.1089/106652700750050862>
- [5] S. Govindarajan and R. A. Goldstein, "On the Thermodynamic Hypothesis of Protein Folding," *Proceedings of the National Academy of Sciences of USA*, Vol. 95, 1998, pp. 5545-5549. <http://dx.doi.org/10.1073/pnas.95.10.5545>
- [6] B. Rates, L. P. Silva, I. C. Ireno, F. S. Leite, M. H. Borges, C. Bloch Jr., M. E. De Lima and A. M. Pimenta, "Peptidomic Dissection of the Skin Secretion of *Phasmahyla jandaia* (Bokermann and Sazima, 1978) (Anura, Hylidae, Phyllomedusinae)," *Toxicon*, Vol. 57, 2011, pp. 35-52. <http://dx.doi.org/10.1016/j.toxicon.2010.09.010>
- [7] The UniProt Consortium, "The Universal Protein Resource (UniProt) in 2010," *Nucleic Acids Research*, Vol. 38, 2010, pp. D142-D148. <http://dx.doi.org/10.1093/nar/gkp846>
- [8] 2012. <http://www.sigmaaldrich.com/life-science/metabolomics/learning-center/amino-acid-reference-chart.html>
- [9] 2011. <http://www.uniprot.org/uniprot/P86636>
- [10] R. Gössler-Schöfberger, G. Hesser, M. Muik, C. Wechselberger and A. Jilek, "An Orphan Dermaseptin from Frog Skin Reversibly Assembles to Amyloid-Like Aggregates in a pH-Dependent Fashion," *FEBS Journal*, Vol. 276, 2009, pp. 5849-5859. <http://dx.doi.org/10.1111/j.1742-4658.2009.07266.x>
- [11] C. Galanth, F. Abbassi, O. Lequin, J. Ayala-Sanmartin, A.

Ladram, P. Nicolas and M. Amiche, "Mechanism of Antibacterial Action of Dermaseptin B2: Interplay between Helix-Hinge-Helix Structure and Membrane Curvature Strain," *Biochemistry*, Vol. 48, 2009, pp. 313-327.  
<http://dx.doi.org/10.1021/bi802025a>

the Elucidation of Membrane-Acting Helical Amphipathic Antimicrobial Peptides," *Current Pharmaceutical Biotechnology*, Vol. 12, 2011, pp. 1184-1193.  
<http://dx.doi.org/10.2174/138920111796117319>

[12] M. Amiche and C. Galanth, "Dermaseptins as Models for