HD5 and HBD1 Variants' Solvation Potential Energy Correlates With Their Antibacterial Activity Against *Escherichia coli*

William F. Porto,[1,2] Diego O. Nolasco,[1,3] Állan S. Pires,[1,2] Gabriel R. Fernandes,[1] Octávio L. Franco,[1,2,4] Sérgio A. Alencar[1]

[1]*Programa De Pós-Graduação Em Ciências Genômicas E Biotecnologia, Universidade Católica De Brasília, Brasília-, DF, Brazil*

[2]*Centro De Análises Proteômicas E Bioquímicas, Pós-Graduação Em Ciências Genômicas E Biotecnologia, Universidade Católica De Brasília, Brasília-, DF, Brazil*

[3]*Curso De Física, Universidade Católica De Brasília, Brasília, DF, Brazil*

[4]*S-Inova Biotech, Pos Graduação em Biotecnologia, Universidade Catolica Dom Bosco, Campo Grande, Campo Grande, Brazil*

Received 27 June 2015; revised 22 October 2015; accepted 2 November 2015
Published online 13 November 2015 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/bip.22763

ABSTRACT:

The structure-activity relationship of defensins is not clear. It is known that point mutations in HD5 and HBD1 could modify their activities; however, these mutations do not seem to alter their three-dimensional structures. Here, applying molecular dynamics simulations, this relationship was studied in depth. There are modifications in flexibility, solvent accessible surface area and radius of gyration, but these properties are not reflected in the activity. Only alterations in the solvation potential energy were correlated to antibacterial activity against Escherichia coli. *Data here reported could lead to a better understanding of structural and functional aspects of* α- *and* β-*defensins. © 2015 Wiley Periodicals, Inc. Biopolymers (Pept Sci) 106: 43–50, 2016.*

Keywords: α-*defensin;* β-*defensin; point mutations; LD_{50}; molecular dynamics*

This article was originally published online as an accepted preprint. The "Published Online" date corresponds to the preprint version. You can request a copy of any preprints from the past two calendar years by emailing the Biopolymers editorial office at biopolymers@wiley.com.

Additional Supporting Information may be found in the online version of this article.

Correspondence to: Octávio L. Franco; e-mail: ocfranco@gmail.com and Sérgio A. Alencar; e-mail: sergiodealencar@gmail.com

INTRODUCTION

Defensins are small antimicrobial peptides present in multicellular organisms, including fungi,[1] plants,[2] and animals.[3,4] They are classified into four major groups: CSαβ defensins, α-defensins, β-defensins, and θ-defensins, according to their folding and disulfide arrangement.[4,5] It is important to highlight that the CSαβ defensins have been described in plants, fungi, and invertebrates,[1,2,4] while the other three groups (α-, β- and θ-defensins) have been described only in vertebrates.[3]

α- and β-defensins show a very stable and conserved structural scaffold. Their three-dimensional structures are predominantly β-sheet, being stabilized by three intramolecular disulfide bridges formed by a "hallmark" six-cysteine motif.[6] The topology of the three disulfide bonds in α-defensins is Cys^{I}-Cys^{VI}, Cys^{II}-Cys^{IV}, and Cys^{III}-Cys^{V}, whereas the equivalent linkage arrangement in β-defensins is Cys^{I}-Cys^{V}, Cys^{II}-Cys^{IV}, and Cys^{III}-Cys^{VI}.[3] Although α-defensin disulfide bond topology differs from that of β-defensins, their tertiary structures are similar.[7,8]

The mechanism of antimicrobial activity of α- and β-defensins seems to be related to interactions with the bacterial membrane bilayer, which is heavily populated by negatively charged phospholipids.[6] Therefore, the positive net charge seems to be essential to α- and β-defensins' antibacterial activity. However, changes in hydrophobic amino acids, without changing the charge, could also change the antibacterial activity.[9,10] In fact, the α- and β-defensins' structure-activity relationship is unclear. According to data from human α-defensin 5 (HD5)[9] and human β-defensin 1 (HBD1),[10] single amino acid variations could affect their antimicrobial activity.

However, such variations do not seem to result in significant structural modifications, comparing their crystal structures with the wild-type one.[9,10]

In this context, molecular dynamics simulations could shed some light on this relationship. The evaluation of protein structure at the atomic level during a short time period could lead to a better understanding of structure-activity relationships in α- and β-defensins. This method has been extensively used to investigate the effect of point mutations in several human proteins, including aurora-A kinase,[11,12] Ras-related C3 botulinum toxin substrate 1,[13] aldosterone synthase,[14] p53,[15] angiogenin,[16] protein tyrosine phosphatase 1B,[17] receptor tyrosine kinase KIT,[18] lamin A/C protein,[19] P protein,[20] and human guanylin.[21] This kind of study has helped to elucidate the malfunction of these proteins in association with human diseases.

Here, a long-term molecular dynamics simulation was applied to investigate the structural changes driven by point mutations described in previous works for HD5[9] and HBD1[10] and correlate them with their antimicrobial activity against *Escherichia coli*. Besides, it was hypothesized that the solvation potential energy could give a quantitative answer about the impact of amino acid substitutions on α- and β-defensins' antibacterial activity, since this structural property could give the charge distribution on the protein surface. Therefore, in order to assess the correlation of solvation potential energy and antibacterial activity against *Escherichia coli*, two linear regression models were constructed based on molecular dynamics simulations and previously published experimental antibacterial activity of HD5[9] and HBD1 variants,[10] modeling the antibacterial activity as a function of solvation potential energy. In addition, the impact of these variations on structure motion and flexibility was also analyzed.

MATERIALS AND METHODS

Defensin Selection

Experimental data from median lethal dose (LD_{50}) against *E. coli* were retrieved from Rajabi et al.[9] and Pazgier et al.[10] for HD5 and HBD1, respectively. Initially, sequences with nonproteinogenic amino acid residues were discarded. Then, the LD_{50} values were converted to μM. Further, for noise reduction, three standard deviations were subtracted from the LD_{50} values, and the sequences with negative values were removed. None of the selected variants affect any cysteine residue. The selected variants are listed in Table I.

Evolutionary Conservation Analysis

The ConSurf server is a Bioinformatics tool for estimating the evolutionary conservation of amino acid positions in a protein molecule based on the phylogenetic relations between homologous sequences.[22] Using the HD5 and HBD1 mature peptide sequences obtained from RefSeq[23] as inputs, ConSurf carried out a search for close homologous sequences using CSI-BLAST (three iterations and 0.0001 *e*-value cutoff) against the UNIREF-90 protein database.[24,25] The sequences were then clustered and highly similar sequences removed using CD-HIT.[26]

Molecular Modeling

One hundred molecular models for each variant, without three-dimensional structures solved, were constructed by comparative molecular modeling through MODELLER 9.10,[27] using the wild-type structures of HD5 and HBD1 (PDB IDs: 1ZMP[7] and 1KJ5,[8] respectively). The models were constructed using the default methods of automodel and environ classes from MODELLER. The final models were selected according to the discrete optimized protein energy score (DOPE score). This score assesses the energy of the model and indicates the best probable structures. The best models were evaluated through PROSA II[28] and PROCHECK.[29] PROCHECK checks the stereochemical quality of a protein structure through the Ramachandran plot, where good quality models are expected to have more than 90% of amino acid residues in most favored and additional allowed regions, while PROSA II indicates the fold quality. Structure visualization was done in PyMOL (http://www.pymol.org).

Molecular Dynamics Simulations

The molecular dynamics simulations of the ensembles (native wild-type and variants) were carried out in a water environment, using the Single Point Charge water model.[30] The analyses were performed by using the GROMOS96 43A1 force field and the GROMACS 4 computational package.[31] The dynamics used the HD5 and HBD1 tridimensional structures or the three-dimensional models from their respective variants as initial structures, immersed in water inside cubic

Table I Selected Variants of HD5 and HBD1 Used in This Study and Their Respective LD_{50} Values Against *E. coli*

Defensin	Variant[a]	Observed LD_{50} (μM)[b]
HD5	Wild-Type	1.20
	Y27A	1.30
	S17A	0.78
	R09A	3.00
	L29F	1.00
	L29A	1.80
	I22A	0.80
	E21I	0.66
	E21A	0.62
HBD1	Wild-Type	2.77
	S07A	4.52
	Q24A	4.49
	L13K	2.36
	K22A	6.68
	G10A	4.61
	D01K	0.66
	D01A	2.86

[a] The amino acid positions of HD5 and HBD1 are relative to PDB IDs 1ZMP and 1KJ5, respectively.

[b] The LD_{50} values were taken from Rajabi et al.[9] and Pazgier et al.[10]

boxes with a minimum distance of 0.7 nm between peptides and the edges of boxes. This distance provides the smallest possible box without generating artifacts in the simulation; and requires lower computational time, since less bulk solvent molecules are simulated. In addition, Souza and Ornstein[32] demonstrated that a larger box size (1.0 or 1.5 nm) does not afford any extra benefit over a smaller box. Chlorine ions were also inserted at the complexes with positive charges in order to neutralize the system charge. Geometry of water molecules was constrained by using the SETTLE algorithm.[33] All atom bond lengths were linked by using the LINCS algorithm.[34] Electrostatic corrections were made by Particle Mesh Ewald algorithm,[35] with a cut-off radius of 1.4 nm in order to minimize the computational time. The same cut-off radius was also used for van der Waals interactions. The list of neighbors of each atom was updated every 20 simulation steps of 2 fs. The system underwent an energy minimization using 50,000 steps of the steepest descent algorithm. After that, the system temperature was normalized to 310 K for 100 ps, using the velocity rescaling thermostat (NVT ensemble). Then the system pressure was normalized to 1 bar for 100 ps, using the Parrinello-Rahman barostat (NPT ensemble). The systems with minimized energy, balanced temperature and pressure were simulated for 100 ns by using the leap-frog algorithm. Each simulation was repeated three times.

Analyses of Molecular Dynamics Trajectories

Molecular dynamics simulations were analyzed by means of the backbone root mean square deviation (RMSD), by comparing the structures at 0 ns and 100 ns, radius of gyration (Rg) and solvent accessible surface area (SASA) using the g_gyrate and g_sas built-in functions of the GROMACS package,[31] respectively. Rg and SASA were plotted as boxplots, since this allows the visualization of the fluctuation of a given property and also indicates the range in which at least 50% of data lies. The covariance matrices of wild-type and variant peptides were constructed using the main chain atoms. The traces of covariance matrix were measured using the g_covar utility of the GROMACS package.

Solvation Potential Energy Calculation

One snapshot from each trajectory was taken at 100 ns for solvation potential energy calculation. The snapshots were taken using the utility trjconv from GROMACS. The conversion of pdb files into pqr files was perfomed by the utility PDB2PQR using the AMBER force field.[36] The grid dimensions for APBS calculation were also determined by PDB2PQR. Solvation potential energy was calculated by APBS.[37]

Correlation with LD_{50} Calculation

The correlation between the LD_{50} and the values of trace covariance matrix and average values of SASA, Rg and solvation potential energy was measured by R^2, using the R package for statistical computing (http://www.r-project.org). For SASA and Rg, the average of 2501 points from simulation 1 was used; for trace of covariance matrix, the average of the three simulations was used; and for solvation potential energy, the average of three structures at 100 ns was used.

Regression Model Construction for Prediction of Median Lethal Dose against *Escherichia coli*

For each simulation replica, the solvation potential energy was measured as described above. The LD_{50} was modeled as a function of average solvation potential energy by linear regression models built with SVMLight.[38] The models were evaluated through R^2, F-test and root mean squared error (RMSE). For additional validations, three other sets of solvation potential energies were generated from molecular dynamics simulations: (i) 20 ns in triplicate; (ii) 30, 60 and 90 ns from a single simulation; and (iii) three random times from a single simulation. All statistics were calculated using the R package for statistical computing (http://www.r-project.org).

RESULTS AND DISCUSSION

It is known that the charge and hydrophobicity are important to the antimicrobial activity of α- and β-defensins;[9,10] however, their structure-activity relationships are not clear.[9,10] Here, in order to shed some light on these relationships, variants of HD5 and HBD1 were selected from two previous works,[9,10] in order to evaluate the relationships between structural parameters and their activities. Initially, we used a stringent filter to select the best candidates for construction of our linear regression, in order to reduce the noise in data, resulting in a total of nine and eight sequences for the α- and β-defensins (Table I), respectively. Then, ConSurf[22] was used to investigate the conservation patterns of the HD5 and HBD1 mature peptide sequences. ConSurf exploits evolutionary variation in multiple sequence alignments, in order to determine the degrees of conservation. The results from this analysis showed that in HD5 all variants studied occur in sites classified as "variable" in terms of conservation, whereas HBD1 showed higher levels of conservation, where two variants occur in sites classified as "variable" (S07A and Q24A), two in "average" (L13K and K22A) and three in "conserved" (D01K, D01A and G10A) (Figure 1). However, there does not seem to be any clear association between levels of conservation and antimicrobial activity, as there are variants that occur in conserved sites and have high activity (HBD1 D01K) and also there are variants that occur in variable sites and have low activity (HD5 Y27A, HD5 R09A, HD5 L29A, HBD1 S07A, and HBD1 Q24A). Finally, the three-dimensional structures were obtained from the Protein Data Bank or constructed by comparative modeling (Supporting Information Table S1), and then evaluated by molecular dynamics.

The point mutations, in fact, caused structural changes in the α- and β-defensin protein structures at varying levels. However, the relationships between structure and antimicrobial activity are unclear. SASA, Rg, and trace of covariance matrices (flexibility) seem to have little influence in α- and β-defensins' antimicrobial activity (Table III).

Taking into account the HD5 variants I22A, Y27A and L29A, they showed a lower SASA than wild-type (Figure 2). Nevertheless, comparing their antimicrobial potency, they showed higher, similar and lower levels compared to wild-

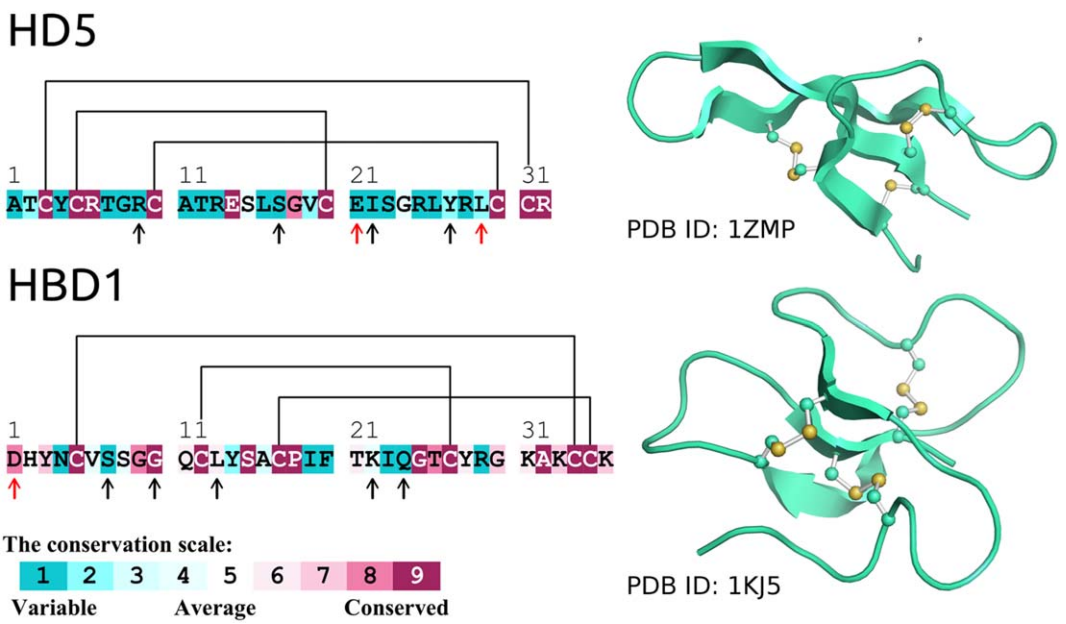

FIGURE 1 Conservation pattern of amino acid residues in HD5 and HBD1 obtained from multiple sequence alignment using ConSurf. The HD5 and HBD1 amino acids are colored based on their conservation grades and conservation levels, where color intensity increases with degree of conservation. A grade of 1 indicates rapidly evolving (variable) sites, which are color-coded in turquoise; 5 indicates sites that are evolving at an average rate, which are colored white; and 9 indicates slowly evolving (evolutionarily conserved) sites, which are color-coded in maroon. The black lines above the sequences indicate the disulfide bridges. The black arrows point to the peptide positions of the variations studied; the red arrows point peptide positions with two variations. The structures of HD5 and HBD1 are shown in cartoon, with disulfide bridges in ball and stick.

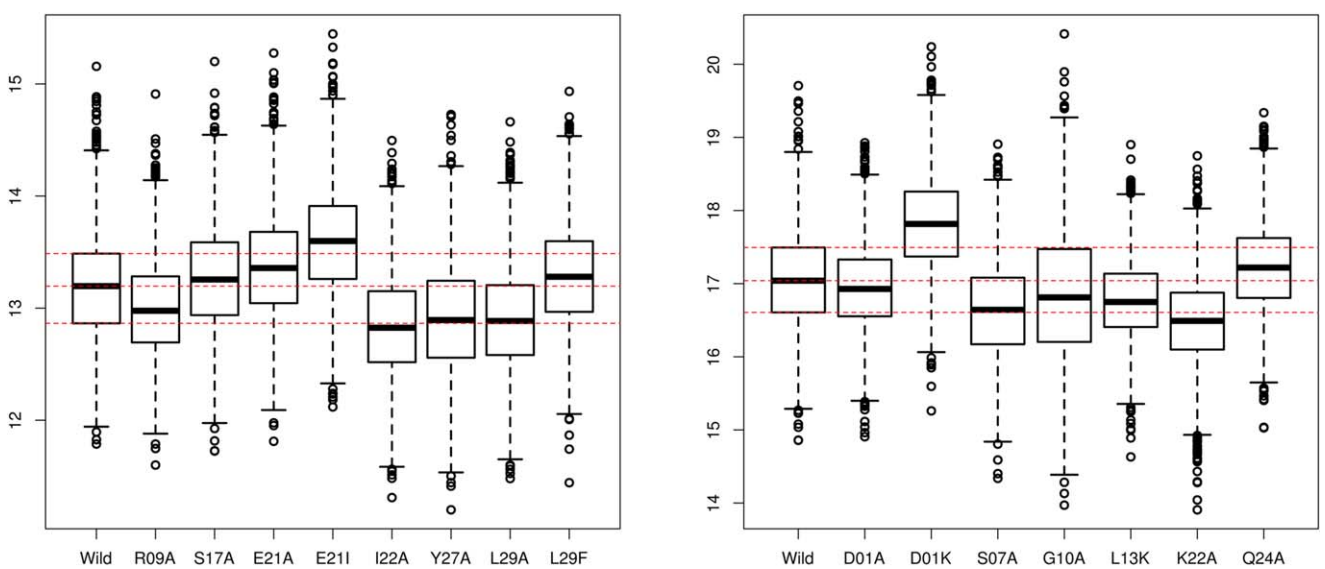

FIGURE 2 The solvent accessible surface area variation of HD5 (left) and HBD1 (right) variants. Dotted red lines indicate the reference values of wild-type defensins. Values are in nm^2.

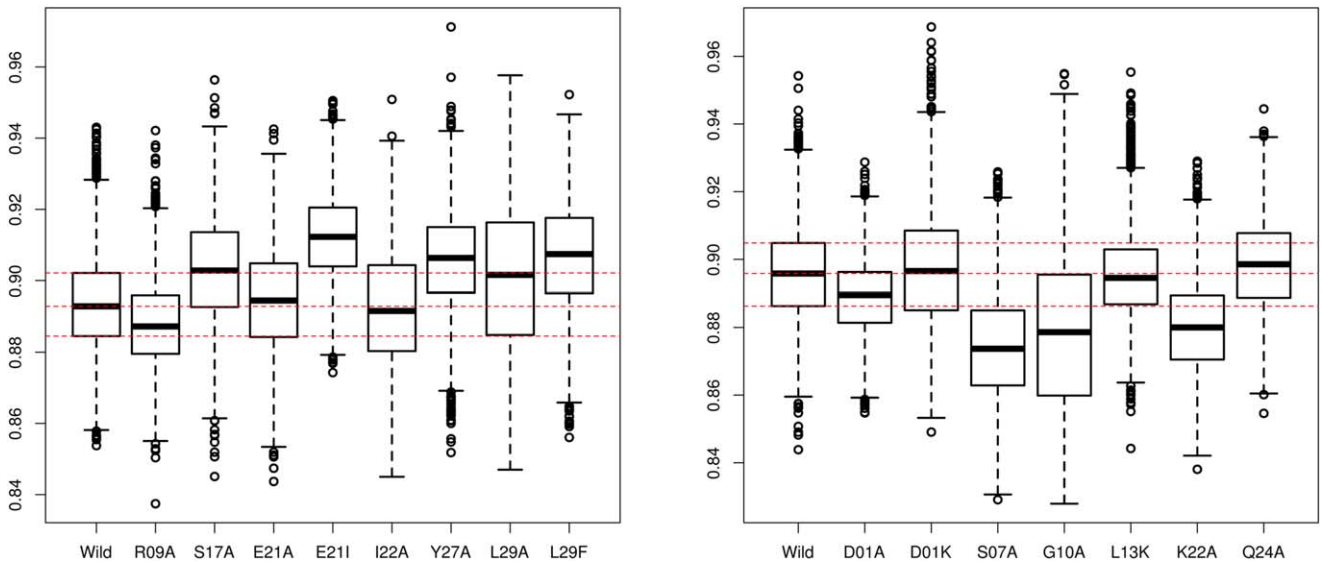

FIGURE 3 The radius of gyration variation during the simulations of HD5 (left) and HBD1 (right) variants. Dotted red lines indicate the reference values of wild-type defensins. Values are in nm.

type, respectively (Table I). Similar data could be observed from SASA of HBD1 variants G10A, D01A, and L13K, which showed similar SASA (Figure 2), but higher, similar and lower levels of antimicrobial potency, respectively, when compared to the wild-type one (Table I). Concerning the Rg, the HD5 variants S17A, L29A, and L29F, showed higher values of Rg, indicating that they have a higher propensity to expand than the wild-type (Figure 3); however, their activities are higher, lower and similar to wild-type, respectively. For HBD1, the variants D01A, D01K and Q24A showed similar Rg values (Figure 3), but also different activities (Table I). In terms of flexibility, a similar situation is observed, where HD5 variants E21I, Y27A and L29A showed higher values of trace of covariance matrix (Table II), but different activities (Table I). As well as for HBD1 variants D01A, L13K and Q24A, which show lower values of trace of covariance

Table II Summary of Molecular Dynamics Results of HD5 and HBD1 Used for Construction of Linear Regression Models

		Trace of Covariance Matrix (nm^2)			RMSD (Å)[b]		
Defensin	Variant[a]	Simulation 1	Simulation 2	Simulation 3	Simulation 1	Simulation 2	Simulation 3
HD5	Wild-Type	1,45703	1,49683	1,57016	2.17	1.37	2.37
	Y27A	1,91382	1,57707	2,03788	1.08	1.30	2.42
	S17A	2,13506	2,33762	1,68029	2.16	2.07	2.00
	R09A	1,27802	1,41801	1,79825	2.61	1.20	2.29
	L29F	1,53505	1,45357	1,50474	2.01	1.47	1.61
	L29A	2,36377	2,39767	3,01148	2.40	1.54	2.42
	I22A	2,30558	1,4542	1,70388	1.27	2.05	2.06
	E21I	1,83656	2,03623	1,59233	1.33	2.41	1.60
	E21A	2,00504	2,22981	2,12639	1.90	2.03	1.33
HBD1	Wild-Type	3,9775	6,12291	4,89279	2.49	2.28	3.18
	S07A	4,36606	1,23891	2,03156	2.94	2.27	2.47
	Q24A	3,05618	3,05618	3,05618	4.27	2.31	1.56
	L13K	1,94932	1,6141	2,61658	1.95	2.13	2.15
	K22A	3,61472	2,2892	6,032	2.02	2.32	3.59
	G10A	8,92232	8,15092	5,68925	2.86	5.19	3.41
	D01K	5,12275	5,79701	5,89107	3.34	3.10	3.36
	D01A	1,94129	4,66188	11,754	2.08	2.87	3.60

[a] The amino acid positions of HD5 and HBD1 are relative to PDB IDs 1ZMP and 1KJ5, respectively.
[b] Data generated by comparing the structure at 0 and 100 ns.

Table III Correlations Between Structural Properties and LD_{50}

	Pearson Correlation Coefficient	
Property	HD5	HBD1
SASA[a]	0.4949	0.7637
Radius of Gyration[a]	0.4935	0.6592
Trace of covariance matrix[b]	−0.1553	−0.1380
Solvation potential energy[c]	0.8536	0.9075

[a] Data generated taking into account 2501 time steps from simulation 1.
[b] Data generated taking into account the average of traces covariance matrix of three simulations.
[c] Data generated taking into account the final structure at 100 ns of three simulations.

matrix (Table II), and also different levels of antimicrobial activities (Table I). The examples above give a small part of the picture, showing the lack of a relationship between these parameters and antimicrobial activity. Although these properties are the most used properties for the evaluation of the impact of amino acid substitutions in several proteins,[11–21] for α- and β-defensins, they seem to have little or no correlation with antimicrobial activity (Table III). The unique property analyzed, which showed correlation to the antimicrobial activity was the solvation potential energy (Table III). Such characteristic seems to be able to reflect the α- and β-defensins mechanism of action, probably being related to bacterial membrane bilayer interactions.[6] Nevertheless, it is important to remark that the exact mechanism of HBD1 has not been precisely established until now. Currently, there are two proposed models for β-defensins interactions with membrane including the (i) formation of channels or (ii) detergent activities.[39] In both cases, the solvation potential energy is important due to the interactions of defensins to negatively charged phospholipids. On the other hand, in the case of HD5, the mechanism of action against *E. coli* has been very well elucidated. It was demonstrated that HD5 internalizes into bacterial cells;[40] further binds to bacterial DNA,[41] which clearly suggests the importance of solvation potential energy in the interaction with the negatively charged phosphate groups in the DNA.

In order to evaluate the antibacterial activity, regression models were constructed for α- and β-defensins, modeling the LD_{50} as a function of average solvation potential energy. As shown in Figure 4, both linear regression models have a good assessment according to the R^2 values (0.7353 and 0.8236 for HD5 and HBD1, respectively), both models showing significant *P*-values (0.0031 and 0.0018 for α- and β-defensins, respectively), using the solvation potential energy data of each replica at 100 ns (Supporting Information Table S2), with RMSE values of 0.368 and 0.721, respectively. The predicted

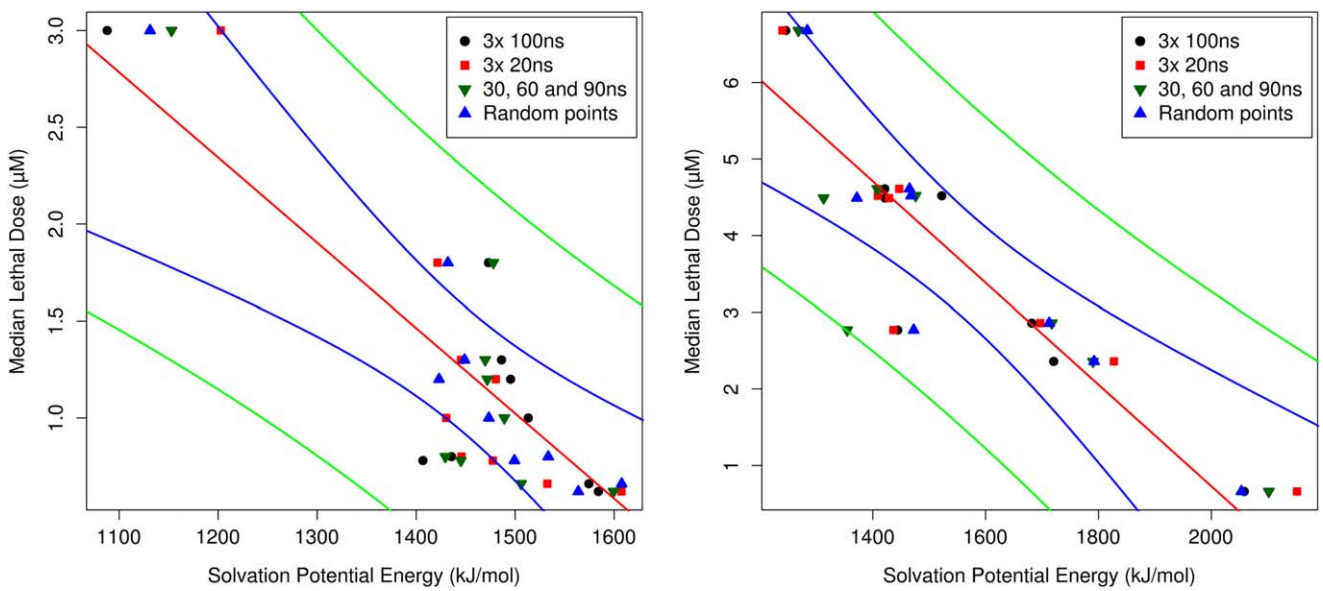

FIGURE 4 Linear regression analysis between solvation potential energy and LD_{50}. Models for α- and β-defensins are represented by the left and right panels, respectively. Black circles represent the data used for model construction (average of solvation potential energy at 100 ns in triplicate). Red squares, green tip down and blue tip up triangles represent the data for model validation, generated by 20 ns in triplicate, three defined points from one simulation (30, 60, and 90 ns) and three random points from one simulation, respectively. The red, blue and green lines represent the linear regression, the confidence (95%) and prediction intervals, respectively.

LD_{50} is described as a function of solvation potential energy by Eqs. (1) and (2) for HD5 and HBD1, respectively:

$$pLD50 = -4.402 esolv \times 10^{-3} + 7.626315 \quad (1)$$

$$pLD50 = -6.634 esolv \times 10^{-3} + 13.998061 \quad (2)$$

In order to assess regression models, data from other points were used including the points from each replica at 20 ns, three defined points (30, 60, and 90 ns) or three random points from a single simulation (Figure 4). These other points have similar assessments, indicating that the models could be used for short simulations and they are in the confidence interval of 95% (Figure 4). It is important to highlight that the relationship between LD_{50} and solvation potential energy seems to be linear, considering the intervals 1087.9 to 1584.1 and 1245.8 to 2058.4 $kJ \cdot mol^{-1}$, for HD5 and HBD1, respectively. Therefore, the correlation could obey a non-linear regression. Although polynomial and radial regression models were also tested, they had worse assessments than the linear one (data not shown).

Antimicrobial activity prediction has gained attention in recent years, and several qualitative methods have been developed.[42–44] Melo et al.[45] developed a quantitative prediction method, which requires *in vitro* data from membrane interactions given by fluorescence spectroscopy assays. However, due to the complexity of antimicrobial peptides' (AMPs) structure and multi-functionality, there are no computational quantitative predictors currently developed. Here, using LD_{50} data from previous studies, we constructed a full computational quantitative prediction method for α- and β-defensins' antibacterial activity based on linear regression using molecular dynamics simulations from molecular models. However, it is important to highlight that there are some limitations regarding our prediction models: (I) There is some degree of uncertainty in the structures of defensin variants. Despite the identities of 96.9 and 97.2% for HD5 and HBD1 variants, the molecular modeling depends on some refinements which could add some noise in the raw data. (II) Since the regression models were constructed based on data from single amino acid mutations of HD5 and HBD1, they may not be applied to other defensins, even from those from the same class (*e.g.*, HD1, HD2, HBD2, HBD3, or other defensins from other organisms). In addition, (III) the regression models do not take into account variations in the cysteine residues, which could generate some degree of structural variation. However, in a previous report, it was demonstrated that the variant C35S has similar structure and is more active than the native wild-type,[46] indicating that the model could also be used for predicting the activity of those variants. In addition, it has been shown that the reduction of all disulfide bridges from HBD1 does not alter its antibacterial activity.[47] Therefore, for the application of such models, it is necessary only to calculate the solvation potential energy through APBS[37] and then apply Eqs. (1) or (2), for HD5 or HBD1 variants, respectively.

In summary, the molecular dynamics simulations provided a clear insight into conformational changes in α- and β-defensin variants; however, these conformational changes seem not to correlate with α- and β-defensins' antimicrobial activity against *E. coli*. The correlation between solvation potential energy and LD_{50} provides a new approach to investigate variations in α- and β-defensins, making prediction prior to synthesis possible, which could be useful for rational experimental design. In conclusion, data here generated could be essential to reduce the number of variations to be screened in genetic association studies in HD5 and HBD1, taking into account the antimicrobial function. In addition, data here reported could lead to a better understanding of structural and functional aspects of α- and β-defensins.

The authors are grateful to the Center for Scientific Computing (NCC/GridUNESP) of São Paulo State University (UNESP) and Dr. T. Joachims, from the Department of Computer Science of Cornell University, USA. This work was supported by CNPq, CAPES, FAPDF and FUNDECT.

REFERENCES

1. Mygind, P. H.; Fischer, R. L.; Schnorr, K. M.; Hansen, M. T.; Sönksen, C. P.; Ludvigsen, S.; Raventós, D.; Buskov, S.; Christensen, B.; De Maria, L.; Taboureau, O.; Yaver, D.; Elvig-Jørgensen, S. G.; Sørensen, M. V.; Christensen, B. E.; Kjaerulff, S.; Frimodt-Moller, N.; Lehrer, R. I.; Zasloff, M.; Kristensen, H. H. Nature 2005, 437, 975–980.
2. Chen, K. C.; Lin, C. Y.; Chung, M. C.; Kuan, C. C.; Sung, H. Y.; Tsou, S. C. S.; Kuo, C. G.; Chen, C. S. Bot Bull Acad Sin 2002, 43, 251–259.
3. Ganz, T. C R. Biol 2004, 327, 539–549.
4. Porto, W. F.; Fensterseifer, G. M.; Franco, O. L. J Mol Model 2014, 20, 2339
5. Cândido, E. S.; Porto, W. F.; Amaro, D. S.; Viana, J. C.; Dias, S. C.; Franco, O. L. In Science against Microbial Pathogens: Communicating Current Research and Technological Advances; Méndez-Vilas, A., Ed.; Formatex, 2011; pp 951–960.
6. White, S. H.; Wimley, W. C.; Selsted, M. E. Curr Opin Struct Biol 1995, 5, 521–527.
7. Szyk, A.; Wu, Z.; Tucker, K.; Yang, D.; Lu, W.; Lubkowski, J. Protein Sci 2006, 15, 2749–2760.
8. Schibli, D. J.; Hunter, H. N.; Aseyev, V.; Starner, T. D.; Wiencek, J. M.; McCray, P. B.; Tack, B. F.; Vogel, H. J. J Biol Chem 2002, 277, 8279–8289.
9. Rajabi, M.; Ericksen, B.; Wu, X.; de Leeuw, E.; Zhao, L.; Pazgier, M.; Lu, W. J Biol Chem 2012, 287, 21615–21627.
10. Pazgier, M.; Prahl, A.; Hoover, D. M.; Lubkowski, J. J Biol Chem 2007, 282, 1819–1829.

11. Kumar, A.; Rajendran, V.; Sethumadhavan, R.; Purohit, R. Cell Biochem Biophys 2013, 66, 787–796.
12. Kumar, A.; Purohit, R. PLoS Comput Biol 2014, 10, e1003318
13. Kumar, A.; Rajendran, V.; Sethumadhavan, R.; Purohit, R. PLoS One 2013, 8, e77453
14. Jia, M.; Yang, B.; Li, Z.; Shen, H.; Song, X.; Gu, W. PLoS One 2014, 9, e104311
15. Chitrala, K. N.; Yeguvapalli, S. PLoS One 2014, 9, e104242
16. Padhi, A. K.; Jayaram, B.; Gomes, J. Sci Rep 2013, 3, 1225
17. Liu, M.; Wang, L.; Sun, X.; Zhao, X. Sci Rep 2014, 4, 5095
18. Purohit, R. J Biomol Struct Dyn 2014, 32, 1033–1046.
19. Rajendran, V.; Purohit, R.; Sethumadhavan, R. Amino Acids 2012, 43, 603–615.
20. Kamaraj, B.; Purohit, R. Cell Biochem Biophys 2013, 68, 97–109.
21. Porto, W. F.; Franco, O. L.; Alencar, S. A. Peptides 2015, 69, 92–102.
22. Ashkenazy, H.; Erez, E.; Martz, E.; Pupko, T.; Ben-Tal, N. Nucleic Acids Res 2010, 38, W529–W533.
23. Pruitt, K. D.; Brown, G. R.; Hiatt, S. M.; Thibaud-Nissen, F.; Astashyn, A.; Ermolaeva, O.; Farrell, C. M.; Hart, J.; Landrum, M. J.; McGarvey, K. M.; Murphy, M. R.; O'Leary, N. A.; Pujar, S.; Rajput, B.; Rangwala, S. H.; Riddick, L. D.; Shkeda, A.; Sun, H.; Tamez, P.; Tully, R. E.; Wallin, C.; Webb, D.; Weber, J.; Wu, W.; DiCuccio, M.; Kitts, P.; Maglott, D. R.; Murphy, T. D.; Ostell, J. M. Nucleic Acids Res 2014, 42, D756–D763.
24. Angermüller, C.; Biegert, A.; Söding, J. Bioinformatics 2012, 28, 3240–3247.
25. Suzek, B. E.; Huang, H.; McGarvey, P.; Mazumder, R.; Wu, C. H. Bioinformatics 2007, 23, 1282–1288.
26. Li, W.; Godzik, A. Bioinformatics 2006, 22, 1658–1659.
27. Webb, B.; Sali, A. Curr Protoc Bioinformatics 2014, 47, 561–565.6.32.
28. Wiederstein, M.; Sippl, M. J Nucleic Acids Res 2007, 35, W407–W410.
29. Laskowski, R.; Macarthur, M.; Moss, D.; Thornton, J. J Appl Cryst 1993, 26, 283–291.
30. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. In Intermolecular Force; Pullman, B., Ed.; Dordrecht, Reidel, 1981; pp 331–342.
31. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. J Chem Theory Comput 2008, 4, 435–447.
32. de Souza, O. N.; Ornstein, R. L. Biophys J 1997, 72, 2395–2397.
33. Miyamoto, S.; Kollman, P. A. J Comput Chem 1992, 13, 952–962.
34. Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. J Comput Chem 1997, 18, 1463–1472.
35. Darden, T.; York, D.; Pedersen, L. J Chem Phys 1993, 98, 10089
36. Dolinsky, T. J.; Nielsen, J. E.; McCammon, J. A.; Baker, N. A. Nucleic Acids Res 2004, 32, W665–W667.
37. Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A. Proc Natl Acad Sci USA 2001, 98, 10037–10041.
38. Joachims, T. In Advances in Kernel Methods - Support Vector Learning; Schölkopf, B., Burges, C., Smola, A., Eds.; MIT-Press, 1999; pp 41–56.
39. Pazgier, M.; Hoover, D. M.; Yang, D.; Lu, W.; Lubkowski, J. Cell Mol Life Sci 2006, 63, 1294–1313.
40. Chileveru, H. R.; Lim, S. A.; Chairatana, P.; Wommack, A. J.; Chiang, I. L.; Nolan, E. M. Biochemistry 2015, 54, 1767–1777.
41. Mathew, B.; Nagaraj, R. Peptides 2015, 71, 128–140.
42. Porto, W. F.; Pires, ÁS.; Franco, O. L. PLoS One 2012, 7, e51444
43. Lata, S.; Sharma, B. K.; Raghava, G. P. S. BMC Bioinformatics 2007, 8, 263
44. Thomas, S.; Karnik, S.; Barai, R. S.; Jayaraman, V. K.; Idicula-Thomas, S. Nucleic Acids Res 2010, 38, D774–D780.
45. Melo, M. N.; Ferre, R.; Feliu, L.; Bardají, E.; Planas, M.; Castanho, M. A. R. B. PLoS One 2011, 6, e28549
46. Circo, R.; Skerlavaj, B.; Gennaro, R.; Amoroso, A.; Zanetti, M. Biochem Biophys Res Commun 2002, 293, 586–592.
47. Schroeder, B. O.; Wu, Z.; Nuding, S.; Groscurth, S.; Marcinowski, M.; Beisner, J.; Buchner, J.; Schaller, M.; Stange, E. F.; Wehkamp, J. Nature 2011, 469, 419–423.