

# Cálculo de Estrutura de Proteínas via Álgebra Geométrica Conforme

**Valter S. Camargo**

Colegiado de Matemática,  
Universidade Estadual do Paraná - Campus Paranavaí  
Av. Gabriel Esperidião, S/nº, Caixa Postal: 306, CEP: 87703-000, Paranavaí - PR  
E-mail: vsc@ime.unicamp.br

**Emerson V. Castelani**

Departamento de Matemática  
Universidade Estadual de Maringá  
Av. Colombo, 5790, CEP 87020-900, Maringá - PR  
E-mail: evcastelani@uem.br

**Carlile Lavor**

Instituto de Matemática, Estatística e Computação Científica - IMECC  
Universidade Estadual de Campinas  
Rua Sergio Buarque de Holanda, 651, CEP: 13083-859, Campinas - SP  
E-mail: clavor@ime.unicamp.br

## RESUMO

Em biologia computacional, um dos problemas mais importantes é a determinação da estrutura tridimensional de uma proteína. Esta estrutura pode ser determinada experimentalmente de duas maneiras: Através de técnicas de Ressonância Magnética Nuclear (RMN) ou técnicas de Cristalografia Raios X ([1]). Nosso trabalho se baseia em dados da primeira, no entanto, de maneira geral, a RMN fornece apenas um conjunto esparsos de distâncias entre os átomos de uma molécula. Neste caso, o problema é determinar a estrutura tridimensional da proteína usando informação sobre distâncias. Na literatura, este problema é conhecido como *Molecular Distance Geometry Problem - MDGP* ([2]) e é geralmente formulado como um problema de otimização contínua. Entretanto, para uma subclasse do MDGP, que inclui as proteínas, é possível tratá-lo como um problema de otimização combinatória ([5]), através de uma formulação discreta, conhecida como *Discretizable Molecular Distance Geometry Problem - DMDGP* ([4]).

O trabalho tem como objetivo apresentar uma nova abordagem ao problema, distinta da conhecida na literatura, para a versão discreta. A metodologia baseia-se na identificação do modelo euclidiano em  $\mathbb{R}^3$ , que envolve o DMDGP, com o Modelo Conforme - MC ([7]) do espaço de Minkowski  $\mathbb{R}^{4,1}$  (ou, Espaço Conforme - EC [6]). Os objetos geométricos envolvidos na formulação do problema se apresentam como primitivas computacionais, o que permitiu a implementação de um novo algoritmo que resolve o DMDGP. O algoritmo proposto utiliza-se da álgebra geométrica em MC, conhecida como Álgebra Geométrica Conforme - AGC ([3]), e das simetrias inerentes ao problema, para resolvê-lo através de uma busca em uma árvore binária, sem procedimentos de recursividade. Também serão apresentados testes computacionais preliminares.

**Palavras-chave:** *Geometria de Distâncias, Estrutura Tridimensional de Proteínas, Biologia Computacional, Otimização Combinatória, Modelo Conforme, Espaço Conforme, Álgebra Geométrica Conforme.*

## Referências

- [1] Brünger, A. T. e Nilges, M., Computational challenges for macromolecular structure determination by X-ray crystallography and solution NMR-spectroscopy, *Quarterly Reviews of Biophysics*, 26, 49-125, 1993.
- [2] Crippen, G. e Havel, T. *Distance Geometry and Molecular Conformation*. Research Studies Press, New York, 1988.
- [3] Fontijne, D., Dorstand, L. e Mann, S. Geometric algebra for computer science: An object-oriented approach to geometry (the morgan kaufmann series in computer graphics), Elsevier, 2007.
- [4] Lavor, C., Liberti, L. e Maculan, N. e Mucherino, A. *The Discretizable Molecular Distance Geometry Problem, Computational Optimization and Applications* 52, 115 - 146, 2012.
- [5] Liberti, L., Lavor, C., Maculan, N. e Mucherino, A. *Euclidean Distance Geometry and Applications*, SIAM Review 56, 3 - 69, 2014.
- [6] Perwass, C. B. U. e Hildenbrand D. *Aspects of Geometric Algebra in Euclidean, Projective and Conformal Space* (an introductory tutorial), Tech. report, 2004.
- [7] Perwass, C. B. U. *Geometric Algebra with Applications in Engineering*, Springer Series in Geometry and Computing, 2009.