

## PLANO DE TRABALHO

DADOS DO PLANO DE TRABALHO	
<b>Projeto de Pesquisa:</b>	PI113495-2016 - Parametrização de campos de força granulados para interação entre moléculas orgânicas em ambiente hidrofóbico
<b>Orientador:</b>	ELTON JOSE FIGUEIREDO DE CARVALHO
<b>Centro:</b>	UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE
<b>Departamento:</b>	ESCOLA DE CIÊNCIAS E TECNOLOGIA
<b>Tipo de Bolsa:</b>	A DEFINIR
<b>Direcionamento(s) da bolsa:</b>	Iniciação Científica Iniciação Tecnológica
<b>Status do Plano:</b>	APROVADO
<b>Cota:</b>	2018-2019 (01/08/2018 a 31/07/2019)
<b>Editais:</b>	[ECT/UFRN] Edital Nº 01/2018 - Edital de Bolsas de Pesquisa da ECT/UFRN
CORPO DO PLANO DE TRABALHO	
<b>Título</b>	Treinamento e estudo em heurísticas de busca no contexto de parametrização de campos de força granulados para dinâmica Molecular
<b>Introdução e Justificativa</b>	<p>Explorar a conformação de moléculas complexas em solução é um problema formidável, tanto em relação ao número de átomos envolvidos quanto em duração das simulações. O espaço de fase pode ser explorado mais amplamente através de simulações mais longas com a contrapartida de uma menor resolução nas interações e coordenadas atômicas. É possível, por exemplo, reduzir a complexidade das moléculas com estratégias granuladas ou "coarse-grained", que consistem em utilizar grãos ou "beads" para representar conjuntos de átomos. Assim, reduz-se o número de partículas na simulação e se elimina vibrações de alta frequência, em troca de uma resolução menor das coordenadas atômicas e simplificação das interações intermoleculares.</p> <p>Há diversas[1] implementações de campos de forças granulados, que diferem nas técnicas de parametrização e na forma com que os átomos são associados aos grãos, o chamado "mapeamento". Em particular, o campo de forças MARTINI[2] usa grãos para representar, cada um, três a quatro átomos pesados (elementos que não sejam hidrogênio). Isso permite passos de integração até 20 vezes maiores que aqueles de simulações atomísticas (10 fs a 40 fs) ao eliminar vibrações de alta frequência, com a vantagem de ser facilmente implementável em pacotes de dinâmica molecular existentes, como o GROMACS[3].</p> <p>O campo de forças MARTINI foi inicialmente desenvolvido para lipídios em solução aquosa e posteriormente estendido para proteínas, ácidos nucleicos e açúcares, com recentes aplicações em nanopartículas de carbono, como fulerenos[4, 5], grafeno[6, 7] e nanotubos de carbono[8, 9, 10]. Nele, os grãos interagem entre si tanto através de ligações químicas quanto através de interações não ligadas.</p> <p>Por ter sido implementado inicialmente para lidar com sistemas biológicos, em que o balanço entre as interações envolvendo grupos hidrofílicos e hidrofóbicos executa um papel mais importante que as diferenças mais sutis entre diversas estruturas de carbono presentes em solventes orgânicos, é necessária uma investigação mais cuidadosa a fim de se desenvolver uma extensão do campo de forças MARTINI para lidar com essas sutilezas, como a adoção de um número maior de tipos de grãos, com parâmetros de van der Waals refinados para sistemas baseados em carbono, em particular os aromáticos.</p> <p>Ainda assim, encontrar um mapeamento adequado enquanto se minimiza a quantidade de novos parâmetros a se acrescentar no campo de forças MARTINI continua sendo uma tarefa complexa. Há muitas variáveis em jogo e uma busca manual de valores adequados seria tediosa e potencialmente pouco produtiva. Por isso é interessante a aplicação de heurísticas de busca, algoritmos que procuram, de uma forma pragmática, minimizar uma variável (normalmente chamada "fitness": adequação) que depende de maneira não trivial de uma série de variáveis, sem que seja necessário explorar todo o espaço de parâmetros ou depender de análises potencialmente subjetivas.</p> <p>Entre esses métodos, temos especial interesse em algoritmos como: otimização por enxame de partículas (PSO—Particle Swarm Optimization)[11], em que um conjunto de parâmetros a ser testado se comporta como uma partícula com "velocidade" e que acelera em direção ao ponto de melhor "fitness" encontrado pelas demais partículas; algoritmos genéticos, em que cada conjunto de parâmetros se comporta como um indivíduo de uma espécie e os parâmetros são seus "genes" e em gerações consecutivas novos indivíduos surgem por "cruzamento" dos indivíduos com melhor "fitness" da geração anterior, mais mutações aleatórias. Dentre os objetivos deste plano de trabalho há a procura por outras heurísticas de otimização e estudo da literatura relevante, então esta lista pode aumentar no decorrer do projeto.</p>
<b>Objetivos</b>	<p>Pretendemos estudar as heurísticas de busca e otimização, como enxame de partículas, algoritmo genético e outras, utilizando como contexto tanto toy models quanto o mapeamento e parametrização de moléculas hidrofóbicas no campo de forças MARTINI.</p> <p>Entre os objetivos específicos, estão o treinamento do discente nos algoritmos de busca, nas ferramentas necessárias para realização dos objetivos acima, como programação em Python e shell script, operação de clusters de computadores e dos pacotes de dinâmica molecular GROMACS e softwares acessórios, como VMD e pymol.</p>
<b>Metodologia</b>	O discente iniciará seus estudos da linguagem de programação Python produzindo pequenos programas e funções que realizem os passos relevantes para construir

as heurísticas de busca, entre elas: ler e escrever arquivos, calcular uma função "fitness", tanto em funções de teste padrão (distância à origem, função de Rosenbrock, entre outras)[11], quanto em resultados de simulações de dinâmica molecular com o campo de forças MARTINI.

No caso de dinâmica molecular o "fitness" depende de propriedades como a densidade da fase líquida, a entalpia de vaporização e outras características de corpo[12], que servem como referência para modelar os parâmetros referentes à interação da molécula com outras semelhantes a ela. Para ajustar os parâmetros entre a molécula de interesse e outras, uma abordagem possível é calcular a energia livre de hidratação e a energia livre de partição entre água e octanol, que costumam ser acessíveis experimentalmente e servem de base comum de comparação entre diferentes moléculas.

#### Habilidades Adquiridas

- Programação elementar em Python (ou C), shell script (bash)
- Operação de computadores com sistema operacional Linux, individuais, em rede e em cluster
- Domínio de heurística de busca (enxame de partículas, algoritmo genético etc.)
- Operação do pacote de dinâmica molecular GROMACS e aplicativos associados (VMD, PYMOL etc)
- Operação de pacotes de geração de gráficos (gnuplot, matplotlib, pandas)
- Noções de dinâmica molecular e parametrização de campos de força

#### Referências

- [1] Helgi I. Ingólfsson et al. Wiley Interdisciplinary Reviews: Computational Molecular Science, 2013. doi:10.1002/wcms.1169.  
 [2] Siewert J Marrink et al. The journal of physical chemistry B, 11(27):7812–7824, July 2007. doi:10.1021/jp071097f.  
 [3] Sander Pronk et al. Bioinformatics, 29(7): 845–854, 2013. doi:10.1093/bioinformatics/btt055.  
 [4] Jirasak Wong-Ekkabut et al. Nature nanotechnology, 3(6):363–368, June 2008. doi:10.1038/nnano.2008.130.  
 [5] Robert S G D’Rozario et al. Nanotechnology, 20(11):115102, March 2009. doi:10.1088/0957-4484/20/11/115102.  
 [6] Alexey V. Titov, Petr Král, and Ryan Pearson. ACS Nano, 4(1):229–234, 2010. doi:10.1021/nn9015778.  
 [7] Dan Wu and Xiaoning Yang. The Journal of Physical Chemistry B, 116(39):12048–12056, 2012. doi:10.1021/jp3043939.  
 [8] E Jayne Wallace and Mark S P Sansom. Nanotechnology, 20(4):045101, January 2009. doi:10.1088/0957-4484/20/4/045101.  
 [9] Niladri Patra and Petr Král. Journal of the American Chemical Society, 133(16):6146–6149, 2011. doi:10.1021/ja2009778.  
 [10] Hwanky Lee and Hyungsu Kim. The Journal of Physical Chemistry C, 116(16):9327–9333, 2012. doi:10.1021/jp3010663.  
 [11] J. Vesterstrom and R. Thomsen. Proc. of the 2004 Congress on Evolutionary Computation, 2:1980–1987, 2004. doi: 10.1109/CEC.2004.1331139  
 [12] Carl Caleman et al. Journal of Chemical Theory and Computation, (8):61–74, 2012. doi:10.1021/ct200731v.

#### CRONOGRAMA DE ATIVIDADES

Atividade	2018					2019						
	Ago	Set	Out	Nov	Dez	Jan	Fev	Mar	Abr	Mai	Jun	Jul
REVISÃO BIBLIOGRÁFICA	X	X	X	X	X	X	X	X	X	X		
TREINAMENTO EM PYTHON, SHELL SCRIPT, LINUX	X	X	X	X								
ESTUDO DAS HEURÍSTICAS DE BUSCA		X	X	X	X							
PESQUISA POR NOVAS HEURÍSTICAS DE BUSCA				X	X	X						
IMPLEMENTAÇÃO DAS HEURÍSTICAS EM TOY MODELS E FUNÇÕES DE TESTE			X	X	X	X	X					
TREINAMENTO EM DINÂMICA MOLECULAR (GROMACS, MARTINI, VMD, PYMOL)				X	X	X	X	X				
IMPLEMENTAÇÃO DE HEURÍSTICAS DE BUSCA NO MAPEAMENTO E PARAMETRIZAÇÃO DO CAMPO DE FORÇAS						X	X	X	X	X		
SIMULAÇÕES DE PRODUÇÃO DE MAPEAMENTOS E PARÂMETROS MARTINI PARA SOLVENTES ORGÂNICOS								X	X	X	X	
PREPARAÇÃO DE RESUMOS E TRABALHOS PARA APRESENTAÇÃO										X	X	X

#### PARECER CONSULTORES

Data/Hora	Parecer	Usuário
04/06/2018 14:51	O plano de trabalho está inserido num tema de grande relevância em Simulação Molecular, a produção de uma estatística confiável por uma melhor amostragem do espaço configuracional, dentro de uma escala de tempo acessível de acordo com o potencial computacional disponível. O processo de obtenção de uma função fitness é interessantíssimo e deve proporcionar um grande aprendizado para o aluno, mas também pode ser um problema de considerável dificuldade, o que pode tornar o presente plano de trabalho bem ambicioso para um aluno de graduação. A estratégia que será utilizada para encontrar o fitness poderia ser mais detalhada.	(c1012084)

#### HISTÓRICO DO PLANO DE TRABALHO

Data/Hora	Situação	Tipo de Bolsa	Usuário
04/06/2018 14:51	APROVADO	A DEFINIR	GESTOR DE PESQUISA
19/03/2018 20:45	CONCORRENDO A COTA	A DEFINIR	ELTON JOSE FIGUEIREDO DE CARVALHO (eltonfc)