

PLANO DE TRABALHO

DADOS DO PLANO DE TRABALHO	
Projeto de Pesquisa:	PIJ13495-2016 - Parametrização de campos de força granulosos para interação entre moléculas orgânicas em ambiente hidrofóbico
Orientador:	ELTON JOSE FIGUEIREDO DE CARVALHO
Centro:	UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE
Departamento:	ESCOLA DE CIÊNCIAS E TECNOLOGIA
Tipo de Bolsa:	A DEFINIR
Direcionamento(s) da bolsa:	Continuidade de algum plano do ano anterior Iniciação Científica Possui aderência a uma das áreas de tecnologias prioritárias do Ministério da Ciência, Tecnologia, Inovações e Comunicações (MCTIC)
Status do Plano:	APROVADO
Cota:	2021-2022 (PIBIC) (01/09/2021 a 31/08/2022)
Edital:	EDITAL Nº 01/2021 - INICIAÇÃO CIENTÍFICA
CORPO DO PLANO DE TRABALHO	
Título	
FFFit: biblioteca Python para parametrização de campos de força pro aprendizado de máquina	
Introdução e Justificativa	
<p>Este plano de trabalho é a continuação e aprofundamento dos trabalhos realizados anteriormente neste projeto de pesquisa.</p> <p>O cálculo de propriedades químicas, eletrônicas e dinâmicas de materiais avançados é bastante custosa computacionalmente quando se leva em conta as interações de natureza quântica entre elétrons e átomos. A fim de tornar tratáveis esses problemas, utilizamos métodos que abstraem a complexidade do sistema na forma de interações parametrizadas entre elétrons (modelos semiempíricos e de tight-binding) ou entre os átomos representados classicamente como massas pontuais dotadas de carga elétrica e potenciais de interação entre pares, trios e quartetos de partículas (dinâmica molecular clássica).</p> <p>O conjunto de parâmetros de interação na dinâmica molecular, chamado de campo de força, é desenvolvido a partir de dados experimentais e cálculos quânticos mais custosos. Os campos de força atomísticos de uso mais geral, como OPLS/AA (all-atom optimized potential for liquid simulation) [1] e CHARMM General Force Field (CGenFF) [2] e o campo de forças granulado MARTINI foram originalmente desenvolvidos para sistemas biológicos. Simulações de sistemas não biológicos, como nanotubos de carbono em dispositivos eletrônicos [3] ou da interação de moléculas inorgânicas como nanopartículas de metais/óxidos [4, 5] e sistemas biológicos demandam extensões desses campos de força, parametrizadas de forma consistente com os demais termos do campo.</p> <p>Entretanto, encontrar parametrizações e mapeamentos adequados enquanto se minimiza a quantidade de novos parâmetros a se acrescentar ou alterar no campo de forças 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>Nas etapas anteriores deste Projeto de pesquisa, desenvolvemos uma biblioteca na linguagem Python, chamada FFFit e registrada perante o INPT através da AGIR-UFRN, que implementa os métodos de otimização por enxame de partículas (PSO—Particle Swarm Optimization)[6] e algoritmo genético para parametrizar novas moléculas em campos de força existentes.</p>	
Objetivos	
<p>Neste plano de trabalho, propomos implementar e melhorar as interfaces do código FFFit com os programas GROMACS[7], LAMMPS[8] e ASE[9]. Também propomos melhorar a forma com que o código interage com o ambiente de computação em cluster, como os sistemas NPAD-UFRN e SDumont, através da aplicação inteligente da biblioteca de comunicação MPI.</p> <p>Em paralelo, executaremos simulações de validação do algoritmo parametrizando o campo de forças CGenFF[10] para um conjunto de pequenas moléculas orgânicas, tendo como referência um banco de dados[11] de resultados de simulações.</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 LAMMPS, e softwares acessórios, como ASE, VMD e pymol, bem como em sistemas de clusters de computadores, como o NPAD-UFRN e o SDumont.</p>	
Metodologia	
<p>O discente continuará sua colaboração com o desenvolvimento e validação do pacote de ajuste de campo de forças. Entre as atividades estão:</p> <p>- Revisão da literatura[11] e catalogação de propriedades como a</p>	

densidade da fase líquida, a entalpia de vaporização e outras características de corpo, que servem como referência para modelar os parâmetros referentes à interação da molécula com outras semelhantes a ela.

- Implementação dos programas em Python que executam o ajuste do campo de força para cada molécula de referência, fazendo uso da biblioteca FFFit desenvolvida nas etapas anteriores.
- Desenvolvimento de interfaces entre a biblioteca FFFit e os software GROMACS, LAMMPS e ASE.

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 algoritmos de heurística de busca (enxame de partículas, algoritmo genético etc.)
- Operação dos pacotes de dinâmica molecular GROMACS, LAMMPS e aplicativos associados (ASE, 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] Jorgensen, W. L.; Tirado-Rives, J. Proc. Natl. Acad. Sci. U.S.A., 102, 6665–6670, 2005. doi:10.1073/pnas.0408037102
- [2] K. Vanommeslaeghe et al. J. Comput. Chem., 31, 671, 2010. doi:10.1002/jcc.21367
- [3] Widianta Gomulya et al. Advanced Materials, 25, 2948-2956, 2013. doi:10.1002/adma.201300267
- [4] Antônio Manesco et al. Environmental effects on the geometry of ZnO nanoparticles: multiscale-guided Wulff construction In: XVIII Brazil MRS Meeting, 2019
- [5] Carvalho, E.J.F. Multiscale computational insight into ZnO nanoparticle synthesis by glycerol-urea route In: XVII Brazilian MRS Meeting, 2018
- [6] J. Vesterstrom and R. Thomsen. Proc. of the 2004 Congress on Evolutionary Computation, 2:1980-1987, 2004. doi: 10.1109/CEC.2004.1331139
- [7] Mark James Abraham et al. SoftwareX 1–2, 19-25, 2015. doi:10.1016/j.softx.2015.06.001
- [8] S. Plimpton, Fast Parallel Algorithms for Short-Range Molecular Dynamics, J Comp Phys, 117, 1-19 (1995).
- [9] Ask Hjorth Larsen et al 2017 J. Phys.: Condens. Matter 29 273003 doi:10.1088/1361-648X/aa680e
- [10] K. Vanommeslaeghe et al., J. Chem. Inf. Model. 2012, 52, 3144-3154; K. Vanommeslaeghe et al., J. Chem. Inf. Model. 2012, 52, 3155-3168
- [11] Carl Caleman et al. Journal of Chemical Theory and Computation, (8):61–74, 2012. doi:10.1021/ct200731v.

CRONOGRAMA DE ATIVIDADES

Atividade	2021				2022							
	Set	Out	Nov	Dez	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago
REVISÃO DE LITERATURA	X	X	X	X	X	X	X	X	X	X		
TREINAMENTO DO BOLSISTA NAS LINGUAGENS (PYTHON, BASH, C) E NAS FERRAMENTAS (SSH, SLURM, MPI)	X	X										
TREINAMENTO DO BOLSISTA NOS PACOTES DE DINÂMICA MOLECULAR GROMACS, LAMMPS, ASE, PYMOL)		X	X									
APLICAÇÃO DO ALGORITMO NA PARAMETRIZAÇÃO DE CAMPOS DE FORÇA PARA MOLÉCULAS ORGÂNICAS SIMPLES		X	X	X	X							
APLICAÇÃO DO ALGORITMO NA OTIMIZAÇÃO DO CAMPO DE FORÇA CGENFF NA SÍNTESE DE NANOPARTÍCULAS DE ZNO			X	X	X							
APLICAÇÃO DO ALGORITMO NA OTIMIZAÇÃO DE CAMPO DE FORÇA COARSE-GRAINED PARA FOLHAS DE GRAFENO				X	X							
SIMULAÇÕES DE PRODUÇÃO					X	X	X	X	X	X	X	X
PREPARAÇÃO DE ARTIGOS E TRABALHOS PARA CONGRESSOS					X	X	X	X	X	X	X	X

HISTÓRICO DO PLANO DE TRABALHO

Data/Hora	Situação	Tipo de Bolsa	Usuário
21/05/2021 03:01	CONCORRENDO A COTA	A DEFINIR	ELTON JOSE FIGUEIREDO DE CARVALHO (<i>elton.carvalho</i>)
01/06/2021 15:52	DISTRIBUÍDO AUTOMATICAMENTE	A DEFINIR	JEFFERSON FERREIRA DE OLIVEIRA (<i>jefferson.oliveira</i>)