

**Universidade Federal do Rio Grande do Sul
Programa de Pós-Graduação em Ciência dos Materiais**

**USO DE ALGORITMO GENÉTICO E SIMULAÇÕES ATOMÍSTICAS PARA
OTIMIZAÇÃO E ANÁLISE DAS PROPRIEDADES DE NANOFILAMENTOS DE
CARBONO**

Pedro Guerra Demingos

Dissertação de Mestrado

Porto Alegre, Junho de 2021

**Universidade Federal do Rio Grande do Sul
Programa de Pós-Graduação em Ciência dos Materiais**

**USO DE ALGORITMO GENÉTICO E SIMULAÇÕES ATOMÍSTICAS PARA
OTIMIZAÇÃO E ANÁLISE DAS PROPRIEDADES DE NANOFILAMENTOS DE
CARBONO**

Pedro Guerra Demingos

Dissertação realizada sob a orientação da Prof. Naira Maria Balzaretto e co-orientação do Prof. André R. Muniz, apresentada ao Programa de Pós-Graduação em Ciência dos Materiais da Universidade Federal do Rio Grande do Sul para obtenção do título de Mestre em Ciência dos Materiais.

Porto Alegre, Junho de 2021

AGRADECIMENTOS

Agradeço ao professor André pela orientação e por ter me dado a chance de estudar as coisas estranhas que os elétrons fazem.

Agradeço à professora Naira pela atenção e pelo tempo dedicado a responder minhas perguntas sobre física, ainda mais estranhas.

SUMÁRIO

1	INTRODUÇÃO	1
2	REVISÃO BIBLIOGRÁFICA	5
2.1	NANOFILAMENTOS DE CARBONO	5
2.2	MÉTODOS DE SIMULAÇÃO ATOMÍSTICA	10
2.2.1	Métodos de Primeiros Princípios	10
2.2.2	Teoria do Funcional da Densidade	12
2.2.2.1	Teoremas de Hohenberg-Kohn	13
2.2.2.2	Equações de Kohn-Sham	15
2.2.2.3	Funcionais de Troca e Correlação	16
2.2.2.4	Outros aspectos de cálculos DFT	17
2.2.3	Dinâmica Molecular Clássica	18
2.2.3.1	ReaxFF	20
2.3	ALGORITMOS GENÉTICOS	22
2.3.1	Algoritmos genéticos aplicados ao design computacional de materiais	24
2.3.1.1	Genótipos	26
2.3.1.2	Cálculo da Aptidão	27
2.3.1.3	Formação de uma nova geração	30
3	MÉTODOS COMPUTACIONAIS	35
3.1	ANÁLISE DAS PROPRIEDADES DE NANOFILAMENTOS DE CARBONO DERIVADOS DE MOLÉCULAS HETEROCÍCLICAS DE CINCO MEMBROS	35
3.2	DESENVOLVIMENTO DE UM ALGORITMO GENÉTICO PARA EXPLO- RAÇÃO DE ESTRUTURAS COMPLEXAS DE NANOFILAMENTOS DE CARBONO	37
3.2.1	Entradas (<i>inputs</i>)	38
3.2.2	Estratégia de criação dos genótipos	39
3.2.3	Simulação e Cálculo de Aptidão Bruta	45
3.2.4	Cálculo de Aptidão, Similaridade, Singularidade e Probabilidade de seleção dos candidatos	48

3.2.5	Produção de uma nova geração: Replicação, Mutação, Cruzamento e Aleatorização	50
3.2.6	Saídas (<i>outputs</i>)	52
3.2.7	Teste do algoritmo genético	53
4	RESULTADOS E DISCUSSÃO	56
4.1	CÁLCULOS DE DFT DAS PROPRIEDADES DE NANOFILAMENTOS DE CARBONO DERIVADOS DE MOLÉCULAS HETEROCÍCLICAS DE CINCO MEMBROS	56
4.1.1	Estrutura Atômica	56
4.1.2	Estabilidade Relativa	58
4.1.3	Propriedades Eletrônicas	60
4.1.4	Propriedades Mecânicas	65
4.1.5	Nanofilamentos de carbono derivados de fosfol	67
4.2	DESENVOLVIMENTO DE UM ALGORITMO GENÉTICO PARA EXPLORAÇÃO DE ESTRUTURAS COMPLEXAS DE NANOFILAMENTOS DE CARBONO	69
4.2.1	Procura por nanofilamentos de benzeno com alta estabilidade	69
4.2.2	Procura por nanofilamentos de benzeno com alta rigidez	74
4.2.3	Simulação de nanofilamentos de outras moléculas aromáticas	75
5	CONCLUSÕES	78
	REFERÊNCIAS	80

RESUMO

Nanofilamentos de carbono (NTHs) são materiais unidimensionais sintetizados a partir de moléculas aromáticas variadas, como benzeno e compostos heterocíclicos. NTHs despertaram grande interesse científico e tecnológico devido às suas propriedades mecânicas, que combinam a resistência dos nanomateriais de carbono com a flexibilidade dos polímeros. Entretanto, há carências na literatura quanto ao entendimento das relações estrutura-propriedade desses materiais, como a ausência de trabalhos teóricos sobre nanofilamentos derivados de moléculas heterocíclicas de cinco membros, e a dificuldade de se estudar sistematicamente estruturas complexas desses materiais (formadas a partir de benzeno). O presente trabalho procura contribuir em ambos os aspectos a partir de uma combinação de cálculos de Teoria do Funcional da Densidade (DFT), simulações de dinâmica molecular, e a implementação de um algoritmo genético. Cálculos DFT revelaram que NTHs derivados de tiofeno, furano e pirrol possuem boa estabilidade e propriedades mecânicas excelentes, bem como um *band gap* eletrônico inferior ao dos materiais derivados de benzeno, devido à presença de heteroátomos, o que sugere que esses NTHs podem ser usados em aplicações nanoeletrônicas. O algoritmo genético implementado foi capaz de gerar NTHs complexos com diferentes propriedades otimizadas, como estabilidade relativa (baixa energia) e rigidez (alto Módulo de Young), mostrando ser um método eficaz para a exploração sistemática de novas estruturas desses materiais.

ABSTRACT

Carbon nanothreads (NTHs) are one-dimensional materials synthesized from varied aromatic molecules, like benzene and heterocyclic compounds. NTHs arouse great scientific and technological interest due to their mechanical properties, which combine the strength of carbon nanomaterials with the flexibility of polymers. However, there are gaps in the literature in regard to the understanding of the structure-property relationships in these materials, such as the lack of theoretical works on NTHs derived from five-membered heterocyclic molecules, and the difficulty of systematically studying the complex structures of these materials (formed from benzene). This work attempts to contribute to both questions with a combination of Density Functional Theory (DFT) calculations, molecular dynamics simulations, and the implementation of a genetic algorithm. DFT showed that NTHs derived from thiophene, furan and pyrrole exhibit good stability and outstanding mechanical properties, as well as a *band gap* inferior to those of benzene-derived materials, due to the presence of heteroatoms, which suggests that these NTHs may be potentially used in nanoelectronic applications. The implemented genetic algorithm was able to generate complex NTHs with different optimized properties, such as relative stability (low energy) and stiffness (high Young's Modulus), showing to be an effective method for the systematic exploration of new structures for these materials.

LISTA DE ILUSTRAÇÕES

FIGURA 1 – Mecanismo inicialmente proposto para a síntese de NTHs a partir de benzeno cristalino.	5
FIGURA 2 – Estruturas teóricas de NTHs estudadas antes da primeira síntese experimental reportada.	6
FIGURA 3 – Estruturas de NTHs que se adequaram melhor à caracterização por RMN de amostras reais.	8
FIGURA 4 – Estrutura realista de um NTH de acordo com Duan et al., partindo-se de resultados experimentais de NMR.	8
FIGURA 5 – Operações de mutação e cruzamento realizadas em genótipos ilustrativos.	23
FIGURA 6 – Cruzamento realizado em genótipos codificados em base 10 e base 2. . .	27
FIGURA 7 – Simulação evolutiva de vacâncias de S em MoS ₂	28
FIGURA 8 – Nanopartículas <i>core-shell</i> de dois genes, bem como todos os possíveis cruzamentos entre elas.	29
FIGURA 9 – Cruzamento com corte em um e dois pontos.	32
FIGURA 10 – Estruturas de NTHs derivados de furano.	35
FIGURA 11 – Exemplos de NTHs de classe II e III.	37
FIGURA 12 – Exemplo de código genético para um NTH de benzeno.	40
FIGURA 13 – Classificação de ligações em moléculas representativas.	44
FIGURA 14 – Representação completa do genótipo, incluindo índices de rotação. . . .	45
FIGURA 15 – Pré-otimização estrutural de um NTH periódico.	47
FIGURA 16 – Comparação da energia prevista por DFT e pelo ReaxFC para diferentes NTHs simples.	54
FIGURA 17 – Estrutura dos NTHs estudados.	57
FIGURA 18 – Estrutura atômica do N- <i>syn-anti</i> em duas geometrias diferentes: com e sem ligações de hidrogênio.	60
FIGURA 19 – Energia de NTHs representativos durante MD.	60
FIGURA 20 – Propriedades eletrônicas dos NTHs.	62
FIGURA 21 – Mapas de densidade de carga das estruturas <i>syn</i>	63
FIGURA 22 – Intervalos de <i>band gaps</i> de NTHs derivados de diferentes moléculas. . .	64

FIGURA 23 – Propriedades mecânicas dos NTHs.	66
FIGURA 24 – Estruturas e propriedades eletrônicas dos NTHs derivados de fosfol. . . .	68
FIGURA 25 – Estruturas representativas de NTHs de benzeno com 4, 8 e 16 moléculas.	70
FIGURA 26 – Energia média e dos três candidatos mais estáveis em um processo evolu- tivo com NTHs de benzeno com 8 moléculas.	71
FIGURA 27 – Distribuição de energia relativa dos candidatos simulados via processo evolutivo e procura aleatória, com 4, 8 e 16 moléculas.	72
FIGURA 28 – Módulo de Young 1D do candidato mais rígido durante processo evolutivo para otimização da rigidez.	74
FIGURA 29 – Estruturas representativas de NTHs derivados de outras moléculas aromá- ticas.	76
FIGURA 30 – Energia do candidato mais estável em evoluções para diferentes precursores.	77

LISTA DE TABELAS

TABELA 1 – Propriedades mecânicas comparadas de nanomateriais unidimensionais de carbono	7
TABELA 2 – Identificação de ligações cruzadas em moléculas com seis átomos ligantes.	43
TABELA 3 – Identificação de ligações cruzadas em qualquer molécula.	43
TABELA 4 – Propriedades calculadas dos NTHs estudados	59

1 INTRODUÇÃO

Nanofilamentos de carbono (*Carbon Nanothreads*, NTHs) são materiais de grande interesse científico e tecnológico, uma vez que combinam a excelente resistência mecânica dos nanomateriais de carbono (como nanotubos e grafeno) à flexibilidade dos polímeros [1, 2]. NTHs são produzidos a partir da compressão cuidadosa de compostos aromáticos, durante a qual as moléculas se ligam umas às outras ao longo de caminhos unidimensionais, produzindo filamentos ultrafinos com superfície hidrogenada [3, 4]. Já foram reportados NTHs produzidos a partir de diferentes moléculas precursoras, como benzeno [3], piridina [5], anilina [6], furano [7] e tiofeno [8], demonstrando que essa família de materiais apresenta uma ampla gama de estruturas possíveis.

NTHs derivados de moléculas aromáticas de seis átomos (como benzeno, piridina e anilina) foram investigados em diversos estudos computacionais recentes, que utilizaram estruturas simplificadas para representá-los, nas quais o padrão de ligações químicas geradas se repete a cada duas ou quatro moléculas [2, 4, 9]. Os resultados demonstraram que esta classe de materiais possui uma gama de propriedades promissoras, que lhes conferem grande potencial para aplicações tecnológicas. No entanto, a caracterização recente de amostras reais desses materiais mostrou que eles possuem uma estrutura mais complexa, com alternância entre os pequenos blocos investigados até então [10, 11], o que exige um estudo mais abrangente de suas possíveis estruturas e propriedades decorrentes.

Os NTHs sintetizados a partir de moléculas heterocíclicas de cinco membros (como furano e tiofeno), por outro lado, possuem uma estrutura mais periódica, uma vez que as reações que lhes dão origem são limitadas devido ao número menor de átomos de carbono por molécula (quatro, em vez de seis) [7, 8]. Portanto, ao contrário dos NTHs derivados de moléculas de seis membros, os materiais derivados de furano e tiofeno podem ser estudados a partir de células unitárias menores. No entanto, em função de sua síntese recente (2019-2021), até o momento nenhum trabalho computacional reportou propriedades fundamentais de interesse tecnológico desses NTHs (como estrutura de bandas eletrônicas, propriedades mecânicas sob tração, e estabilidade relativa de isômeros).

O presente trabalho tem como objetivo central contribuir para um maior entendimento das estruturas e propriedades de NTHs. Mais especificamente, este estudo visa suprir as duas

carências na literatura mencionadas acima: (i) ausência de estudos computacionais para estimativa de propriedades físicas de NTHs derivados de compostos heterocíclicos de cinco membros, e (ii) a falta de investigações sobre NTHs derivados de benzeno com estruturas mais complexas, nas quais o padrão de ligações químicas varia ao longo de um grande número de moléculas precursoras.

De forma a atender ao item (i), foram conduzidos cálculos baseados na Teoria do Funcional da Densidade (DFT) para a elucidação das propriedades estruturais, mecânicas e eletrônicas de NTHs derivados de furano, tiofeno, pirrol e fosfol. A escolha das duas primeiras moléculas se deve à sua síntese já ter sido reportada [7, 8], e a seleção das duas últimas tem o intuito de encorajar futuros esforços experimentais para a obtenção de novos materiais. Em particular, essa investigação se propõe a elucidar as seguintes questões principais:

- Se esses materiais são estáveis, isto é, se eles têm menor energia do que as moléculas das quais se originam;
- Como a presença e a disposição de heteroátomos afeta a estrutura eletrônica do material, potencialmente alterando seu *band gap* eletrônico e possibilitando novas aplicações;
- Como a presença de diferentes heteroátomos altera as propriedades mecânicas do material, em particular o comportamento tensão-deformação durante testes de tração uniaxial.

Os resultados obtidos na investigação descrita até aqui resultaram no artigo *First-principles study of carbon nanothreads derived from five-membered heterocyclic rings: thiophene, furan and pyrrole*, publicado na *Physical Chemistry Chemical Physics* em 2021 (*Phys. Chem. Chem. Phys.*, 23, 2055) [12].

O problema (ii) é mais complexo. A enumeração de NTHs de benzeno idealizados – isto é, com o padrão de ligações repetindo-se a cada duas moléculas – inclui até 50 estruturas diferentes, das quais 15 possuem boa estabilidade [4]. Esse estudo idealizado permitiu uma enumeração inicial de NTHs possíveis, bem como de suas propriedades. Entretanto, é necessário fazer uma combinação dessas estruturas ideais para modelar o material real, o que envolve um número exponencialmente grande de possibilidades. Logo, é inviável realizar uma exploração exaustiva das possíveis estruturas complexas, sendo necessária uma estratégia metaheurística que permita a análise sistemática e a otimização de NTHs por amostragem.

A solução proposta aqui é a implementação de um algoritmo genético para a busca e otimização de estruturas complexas de NTHs. Algoritmos genéticos empregam conceitos evolutivos (como *seleção*, *mutação* e *cruzamento*) para descobrir e otimizar candidatos em relação a alguma propriedade de interesse (tratada como *aptidão*). Em particular, o presente desenvolvimento se propõe a:

- Implementar um algoritmo genético específico para estudo de NTHs, que possa ser utilizado por terceiros de maneira simples e sistemática;
- Garantir a flexibilidade do algoritmo em relação às preferências do usuário, em particular no que diz respeito ao conjunto de estruturas a serem investigadas (molécula precursora, número de moléculas por candidato, etc), a propriedade de interesse a ser otimizada, e o método de simulação computacional empregado para calculá-la;
- Demonstrar a validade do algoritmo, mostrando que ele apresenta melhor desempenho que uma busca aleatória ao procurar NTHs de baixa energia.

O desenvolvimento desse algoritmo genético permitirá não apenas a descoberta de estruturas estáveis de NTHs geradas a partir de diferentes moléculas e com variados tamanhos, mas também a otimização de propriedades específicas de interesse prático, como resistência à tração e Módulo de Young. A determinação de estruturas com propriedades otimizadas é de grande interesse tecnológico, uma vez que a síntese de NTHs é controlada cineticamente, o que em princípio permitiria a manipulação de suas características estruturais a partir da variação de parâmetros experimentais [3, 13]. Além disso, o algoritmo pode ser utilizado na elucidação das estruturas de NTHs presentes em amostras reais, através da minimização da diferença entre uma dada propriedade estrutural mensurável do material real e do simulado. Essa propriedade pode ser um descritor invariante em relação a operações de simetria, como, por exemplo, a função de distribuição radial (RDF).

Esta dissertação está estruturada da seguinte forma. A revisão bibliográfica aborda NTHs (Seção 2.1); métodos de simulação atômica, como DFT e dinâmica molecular clássica (Seção 2.2); e algoritmos genéticos (Seção 2.3). A descrição da metodologia está dividida em duas partes: uma relativa aos cálculos de DFT de NTHs derivados de moléculas de cinco membros (Seção 3.1), e uma destinada a descrever a implementação do algoritmo genético (Seção 3.2). A mesma divisão é feita nos resultados: primeiro são apresentados os resultados da

análise de NTHs baseada em cálculos DFT (Seção 4.1), e então os testes utilizados para verificar o funcionamento do algoritmo genético (Seção 4.2). Por fim, são apresentadas as conclusões e as perspectivas para trabalhos futuros (Capítulo 5).

2 REVISÃO BIBLIOGRÁFICA

2.1 NANOFILAMENTOS DE CARBONO

Nanofilamentos de carbono (*Carbon Nanotreads*, NTHs), também chamados em trabalhos iniciais de nanofilamentos de diamante (*Diamond Nanotreads*), são materiais unidimensionais formados majoritariamente por átomos de carbono em hibridização sp^3 [3, 13, 14]. Esses materiais foram sintetizados pela primeira vez em 2015 [3], sendo a primeira obtenção de um produto ordenado a partir da compressão e descompressão lenta de cristais de benzeno (apenas materiais amorfos haviam sido obtidos anteriormente empregando processos similares [15]).

NTHs são formados a partir do alinhamento e união ordenada de moléculas aromáticas (benzeno sendo o exemplo arquetípico), tal que os átomos (em geral de carbono) formem ligações covalentes uns com os outros ao longo de uma direção preferencial, mudando sua hibridização de sp^2 para sp^3 , como mostrado na Figura 1. O resultado é um material extremamente fino e com a superfície externa hidrogenada, possuindo uma composição igual à da molécula que lhe dá origem (com fórmula empírica $(CH)_n$ no caso do benzeno).

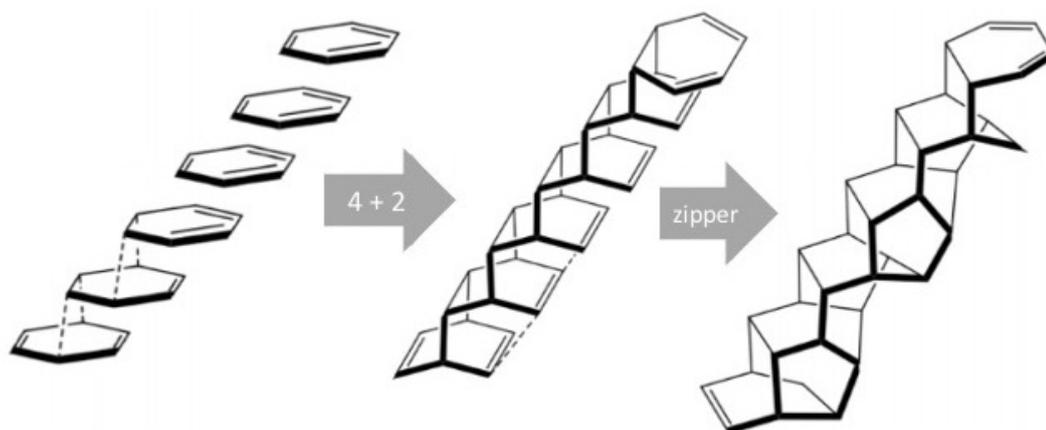


FIGURA 1 – Mecanismo inicialmente proposto para a síntese de NTHs a partir de benzeno cristalino, iniciando com uma cicloadição [4 + 2]. A estrutura final é chamada de *Zipper Polymer*. Adaptado do material suplementar de Fitzgibbons et al. [3].

A síntese mecanoquímica a partir do benzeno forma monocristais aproximadamente hexagonais de NTHs, que se estendem por até centenas de microns, com uma orientação cristalográfica consistente, sem a necessidade de catalisadores ou altas temperaturas [13]. O controle que a tensão aplicada tem sobre a direção da reação é tão significativo que NTHs

orientados são formados não apenas a partir de monocristais de benzeno em fase II (a forma mais estável na pressão de síntese), mas também de sistemas policristalinos multifásicos e em diferentes configurações experimentais. Esses resultados demonstram que a síntese dos NTHs não é topoquímica – isto é, não depende da geometria do cristal de origem –, sugerindo que outras moléculas aromáticas podem ser utilizadas como precursoras, independente de quais sejam suas fases cristalinas.

Em teoria, o material unidimensional formado pode ter estruturas atômicas e morfologias diferentes, dependendo do padrão de ligações carbono-carbono que são formadas durante a reação [4]. Algumas das estruturas mais simples e estáveis de NTHs já haviam sido estudadas computacionalmente antes da primeira síntese, e estão mostradas na Figura 2. A nanoestrutura *Tube(3,0)* havia sido proposta como o nanotubo de carbono com o menor diâmetro possível, cuja superfície externa precisaria ser hidrogenada para garantir estabilidade [16]. A estrutura chamada de *Polymer I* foi encontrada em uma investigação teórica das diversas fases do benzeno cristalino sob altas pressões, tendo se formado espontaneamente em um dos cálculos de otimização estrutural [17]. Por fim, o *Polytwistane* foi sugerido como o produto da polimerização de moléculas de *twistane* [18].

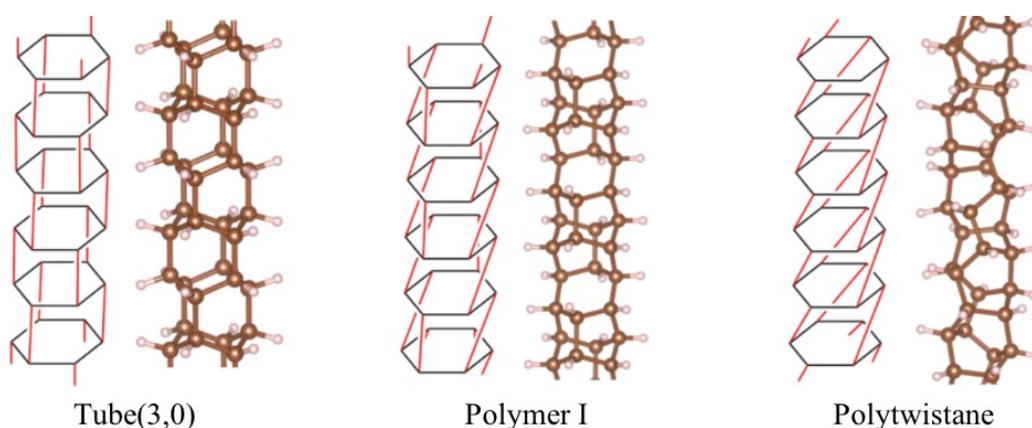


FIGURA 2 – Estruturas teóricas de NTHs estudadas antes da primeira síntese experimental reportada. Reproduzido com permissão de Chen et al. [19], Copyright 2015 American Chemical Society.

Após a primeira síntese de NTHs ser reportada, aumentaram os esforços teóricos e experimentais para a elucidação das possíveis estruturas e propriedades desse novo material. Estudos envolvendo simulações de Dinâmica Molecular clássica e cálculos baseados na Teoria do Funcional da Densidade mostraram que os NTHs possuem excelentes propriedades mecânicas, combinando uma alta resistência à tração (semelhante à do diamante) com a flexibilidade dos

polímeros [1, 2, 20–22]. Em função de sua estrutura predominantemente formada por carbono sp^3 e sua ótima resistência, esse material foi chamado inicialmente de Nanofilamento de Diamante, sendo pensado como o análogo sp^3 de outros materiais unidimensionais de carbono, como nanotubos de carbono (sp^2) e o carbino (sp), cujas propriedades estão mostradas na Tabela 1. Ao contrário do que é observado nestes outros materiais, entretanto, as propriedades mecânicas de NTHs têm baixa sensibilidade a defeitos estruturais, o que os torna uma opção ainda mais atrativa [1].

Material	Força Espec. (10^7 N.m/kg)	Rigidez Espec. (10^8 N.m/kg)
NTHs	1.4 – 2.4	1.0 – 2.0
Nanotubos de carbono	4.0 – 7.5	4.0 – 5.0
Carbino	6.0 – 7.5	10.0

TABELA 1 – Propriedades mecânicas comparadas de nanomateriais unidimensionais de carbono. A força específica é definida como a resistência à tração (tensão no ponto de ruptura) dividida pela massa específica do material, e rigidez específica é o Módulo de Young dividido pela massa específica. Referências: NTHs [2], nanotubos [23], e carbino [24].

Além de suas excelentes propriedades mecânicas, os NTHs possuem um conjunto interessante de propriedades eletrônicas e térmicas [25–32], que levaram à sugestão de diversas aplicações tecnológicas para eles, como a sua utilização em nano ressonadores, nanocompósitos, fibras de carbono, armazenamento de energia, entre outras [33–40].

Em 2018, a caracterização de amostras de NTHs por ressonância magnética nuclear (RMN) em estado sólido revelou a presença de quantidades consideráveis de carbono sp^2 no material [10, 11]. Esses resultados dão suporte à hipótese de que a síntese de NTHs se dá por uma série de cicloadições [4 + 2], no qual as moléculas de benzeno se ligam umas às outras ao longo de uma direção preferencial, saturando 4 de seus 6 carbonos [19]. Acreditava-se até então que os outros 2 átomos reagiriam de maneira a formar o chamado *Zipper Polymer* (como mostrado na Figura 1), mas as caracterizações mais recentes apontam para uma presença significativa do denominado *Square Polymer*, bem como de estruturas parcialmente saturadas com ligações duplas C=C nas laterais da cadeira principal (mostradas na Figura 3). Em função desses resultados, aceita-se atualmente que a melhor representação de um NTH real é um material complexo, formado pela combinação de diversas estruturas ideais (por exemplo, a estrutura ilustrativa da Figura 4). Além disso, acredita-se que a estrutura de NTHs possa ser manipulada através dos parâmetros de síntese, uma vez que a reação de formação desses materiais é controlada cineticamente [3, 4]; a princípio, isso permite o design de NTHs visando a otimização de uma propriedade específica.

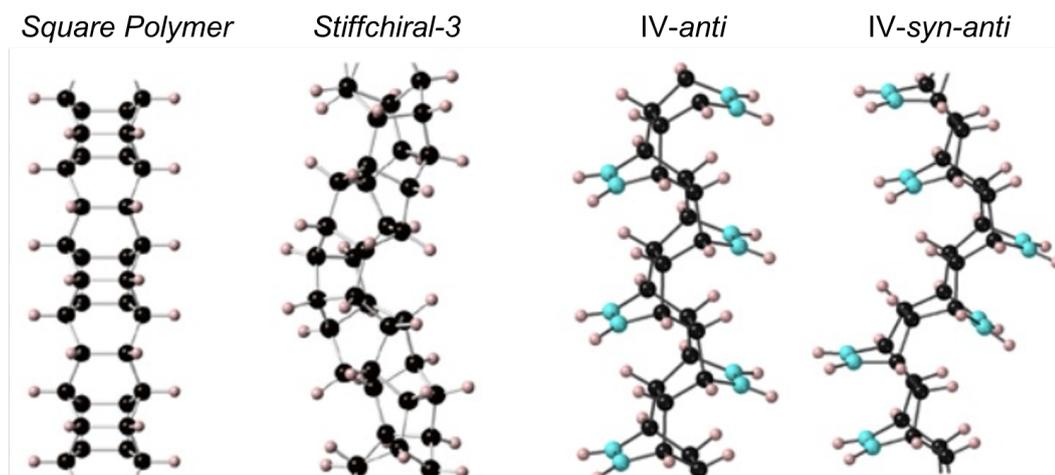


FIGURA 3 – Estruturas de NTHs que se adequaram melhor à caracterização por RMN de amostras reais. Os átomos pretos e azuis representam carbono hibridizado em sp^3 e sp^2 , respectivamente, e os rosas-claros representam hidrogênio. As duas estruturas à direita são parcialmente saturadas, e o "IV" em seu nome indica que possuem quatro (de seis) átomos saturados por molécula. Adaptado com permissão de Wang et al. [11], Copyright 2018 American Chemical Society.

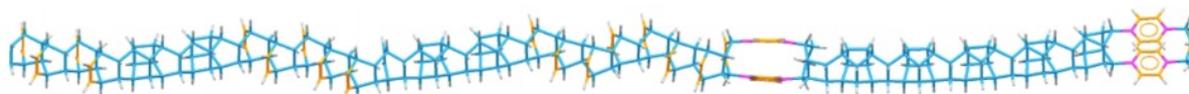


FIGURA 4 – Estrutura realista de um NTH de acordo com Duan et al., partindo-se de resultados experimentais de RMN. Átomos de carbono sp^3 , sp^2 hydrogenados e sp^2 não-hydrogenados estão mostrados em azul, amarelo e magenta, respectivamente. Átomos de hidrogênio estão representados em branco. Reproduzido com permissão de Duan et al. [10], Copyright 2018 American Chemical Society.

Um estudo computacional mostrou que os NTHs parcialmente saturados possuem propriedades mecânicas semelhantes às de suas contrapartes formadas apenas por carbono sp^3 , o que permite que também sejam utilizados em aplicações que exigem resistência, como em fibras de carbono e para o reforço de nanocompósitos [9]. Além disso, a presença de ligações duplas C=C lhes confere um leque maior de propriedades eletrônicas, com *band gaps* variando na faixa de semicondutores a isolantes dependendo da quantidade e orientação de ligações duplas, o que abre uma gama de possíveis novas aplicações, como em dispositivos nanoeletrônicos.

A síntese de NTHs a partir do benzeno impulsionou estudos teóricos de estruturas derivadas de outras moléculas aromáticas. Esses esforços incluíram a investigação de derivados monosubstituídos do benzeno (como anilina, fenol e fluor-benzeno), compostos heterocíclicos (como piridina) [41] e hidrocarbonetos poliaromáticos (como naftaleno, coroneno e muitos outros) [42]. Cálculos computacionais mostraram que a presença de heteroátomos leva a uma maior polaridade nas ligações do material, o que pode contribuir para seu desempenho em

nanocompósitos e sua solubilidade em solventes polares. Além disso, a presença de nitrogênio na cadeia principal mostrou-se capaz de reduzir o *band gap* de NTHs [43–46]. Também foram sugeridas estruturas equivalentes de silício e germânio, que possuem *band gaps* muito reduzidos [47], bem como estruturas bi- e tri-dimensionais derivadas dos NTHs uni-dimensionais [48, 49].

Pouco depois da publicação desses trabalhos teóricos, novas estruturas de NTHs foram sintetizadas a partir de diferentes moléculas aromáticas, como piridina [5], anilina [6, 50] e fenol [51]. O material derivado da piridina também pode ser chamado de Nanofilamento de Nitreto de Carbono, uma vez que possui uma razão molar de N:C de 1:5, como a molécula de origem, o que lhe confere uma quantidade de nitrogênio muito superior à que é possível introduzir a temperatura ambiente em outras nanoestruturas de carbono. A caracterização desse material apontou a presença de ligações duplas C=N, bem como o fenômeno da fotoluminescência [5, 52]. Também foram reportadas sínteses similares a partir de benzeno dissustituído (com diferentes grupos), coranuleno, cubano e adamantano, bem como diversos cocristais (como $C_2H_2 \cdot C_6H_6$, $C_6H_6 \cdot C_6F_6$, e $C_{10}H_8 \cdot C_{10}F_8$, a última possuindo moléculas de naftaleno) [53–60].

Além dos casos citados acima, também foram sintetizados NTHs a partir de moléculas heterocíclicas de cinco membros, como furano e tiofeno [7, 8]. Essas sínteses são de particular interesse uma vez que os heterociclos de cinco átomos possuem apenas quatro carbonos, o que limita a reação de síntese dos NTHs à primeira etapa (a cadeia de cicloadições [4 + 2]). O resultado é um material mais regular ao longo dos nanofilamentos e com maior cristalinidade nas demais dimensões, o que abre a possibilidade de cristais com propriedades anisotrópicas. Esses trabalhos também alcançaram síntese de NTHs em pressões mais baixas, passando dos 20 GPa iniciais para cerca de 10 GPa no caso do furano (graças à sua menor estabilidade por ressonância [7]), o que está próximo dos 6 GPa necessários para produção em escala industrial, como é feito atualmente com o diamante [13].

Em suma, NTHs formam uma família de nanomateriais de carbono com uma enorme gama de possíveis estruturas, uma vez que diversas moléculas aromáticas podem ser utilizadas em sua construção, incluindo cocristais, sem que a morfologia do cristal precursor seja um fator impeditivo (uma vez que a síntese não é topoquímica). Além disso, a síntese a temperaturas relativamente baixas (em torno de 300K) elimina rotas reacionais adicionais, o que lhe confere um alto controle cinético [10, 13], possibilitando um design ainda mais preciso de sua estrutura. Em função dessa flexibilidade estrutural, NTHs podem apresentar uma enorme variedade de propriedades, que os tornam promissores para um grande número de aplicações.

2.2 MÉTODOS DE SIMULAÇÃO ATOMÍSTICA

2.2.1 Métodos de Primeiros Princípios

Uma vez que a predição das propriedades físicas e químicas de sistemas atômicos é de grande interesse científico e tecnológico, diversos métodos de simulação atomística foram desenvolvidos nos últimos cem anos. Os métodos de primeiros princípios, ou *ab initio*, são aqueles que não fazem uso de parâmetros empíricos – ou seja, partem apenas de equações e constantes fundamentais da natureza.

A descrição precisa de sistemas materiais em nível atômico exige um tratamento quântico, isto é, a resolução da equação de Schrodinger (independente do tempo) para as partículas em questão [61]. No entanto, a solução analítica só é possível para sistemas com um único elétron, como é o caso do átomo de hidrogênio. Sistemas com dois ou mais elétrons só podem ser estudados através de métodos numéricos, utilizando-se computadores e um grau maior ou menor de aproximações. A equação de Schrodinger independente do tempo é mostrada na Eq. 2.1, onde E e Ψ são a energia e a função de onda do sistema, e \hat{H} é o seu Hamiltoniano.

$$\hat{H}\Psi = E\Psi \quad (2.1)$$

O Hamiltoniano descreve todas as parcelas de energia do sistema na forma de um operador diferencial. Essas parcelas consistem na energia cinética das partículas individuais e energia potencial decorrente da interação entre partículas (e/ou com um potencial externo). Em sistemas constituídos de elétrons e núcleos atômicos, a única interação relevante é a eletrostática, de maneira que a única parcela de energia potencial se dá pela equação de Coulomb. Portanto, o Hamiltoniano pode ser dividido em cinco parcelas, como mostrado na Equação 2.2, onde os dois primeiros termos são os operadores da energia cinética dos núcleos e dos elétrons, respectivamente; e os três últimos representam a energia potencial entre núcleos, entre elétrons, e entre núcleos e elétrons, respectivamente.

$$\hat{H} = \hat{K}_n + \hat{K}_e + \hat{U}_{nn} + \hat{U}_{ee} + \hat{U}_{en} \quad (2.2)$$

Cada um desses termos pode ser escrito como um operador diferencial, resultando nas parcelas da expressão 2.3, mostradas na mesma ordem da Equação anterior. Os somatórios abrangem os núcleos I/J e os elétrons i/j , cujas massas são M e m , respectivamente, e cujas

posições são dadas pelos vetores \mathbf{R} e \mathbf{r} . Além disso, \hbar é a constante de Planck reduzida, e é a carga fundamental, e Z é a carga dos núcleos.

$$\begin{aligned}
 H = & - \sum_{I=1}^P \frac{\hbar^2}{2M_I} \nabla_I^2 - \sum_{i=1}^N \frac{\hbar^2}{2m_i} \nabla_i^2 \\
 & + \frac{e}{2} \sum_{I=1}^P \sum_{J \neq I}^P \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} + \frac{e}{2} \sum_{i=1}^N \sum_{j \neq i}^N \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} - \frac{e}{2} \sum_{I=1}^P \sum_{i=1}^N \frac{Z_I}{|\mathbf{R}_I - \mathbf{r}_i|}
 \end{aligned} \tag{2.3}$$

Como mencionado anteriormente, essa equação é complexa demais para que uma solução rigorosa seja possível, mesmo utilizando-se recursos computacionais modernos. Portanto, algumas simplificações são necessárias. Uma vez que os núcleos têm uma massa mais de mil vezes superior à dos elétrons, seus movimentos são muito mais lentos, o que possibilita algumas aproximações. A aproximação de Born-Oppenheimer consiste em tratar separadamente as funções de onda dos núcleos e as dos elétrons, e a aproximação adiabática consiste em tratar as funções de onda dos elétrons como se estivessem constantemente no estado fundamental para o potencial externo no qual se encontram (neste caso, aquele produzido pelos núcleos) [62, 63]. Dessa forma, os núcleos seguem a sua própria dinâmica, enquanto os elétrons se adaptam instantaneamente ao potencial gerado por eles; logo, o problema eletrônico é separado do problema nuclear. Além disso, uma aproximação adicional pode ser feita: os núcleos podem ser tratados como partículas clássicas, uma vez que seus efeitos quânticos podem ser desprezados na predição da maior parte das propriedades da matéria. O problema que resta é, portanto, resolver a equação de Schrodinger para os elétrons do sistema, considerando que os núcleos ocupam posições fixas. A expressão resultante está mostrada na Equação 2.4, onde a interação com os núcleos \hat{U}_{en} passa a ser tratado como um potencial fixo (ao qual pode ser somada uma contribuição externa, como um campo elétrico aplicado).

$$\hat{H} = \hat{K}_e + \hat{U}_{ee} + \hat{U}_{en} \tag{2.4}$$

Mesmo com essas aproximações, entretanto, o problema continua complexo, uma vez que segue envolvendo múltiplos elétrons interagindo entre si. Logo, mais simplificações são necessárias, e é neste ponto que a maioria dos métodos quânticos para cálculo da estrutura eletrônica se diferenciam.

O primeiro e mais simples método foi proposto por Hartree em 1928 [64], e consiste

em escrever a função de onda de muitos elétrons Ψ como o produto das funções de onda de cada elétron individual ψ_i . Dessa forma, a função de onda de cada elétron é obtida separadamente, modelando-se a interação com os demais elétrons através de um campo externo efetivo. Trata-se, portanto, de um modelo de campo médio, que ignora efeitos quânticos mais sutis, como a correlação eletrônica (discutida mais adiante). Nesse algoritmo, para se calcular a função de onda de um elétron, é preciso conhecer a dos demais, de maneira que a solução completa deve ser obtida de maneira iterativa, repetindo-se o ciclo até que seja alcançada a auto-consistência.

Uma melhoria pode ser feita no método de Hartree ao escrever a função de onda na forma de uma expansão em determinantes de Slater, que garante que a função de onda do sistema seja anti-simétrica, como exige a natureza fermiônica dos elétrons [65]. Isso considera de maneira precisa a parcela de troca do Hamiltoniano (discutida abaixo). Com esta alteração e a inclusão de termos adicionais de acoplamento nas equações diferenciais, o modelo resultante é chamado de método de Hartree-Fock [66].

Duas parcelas de energia citadas nos últimos parágrafos são relacionadas aos chamados efeitos de troca e correlação. A troca é um efeito da mecânica quântica que ocorre entre partículas idênticas, resultante da necessidade de que a função de onda que as descreve seja invariante com relação à troca entre quaisquer duas dessas partículas (exceto pelo sinal) [61]. Esse efeito produz uma interação adicional entre as partículas, alterando a energia do sistema; no caso dos elétrons, ele também é chamado de Repulsão de Pauli, estando associado ao Princípio da Exclusão de Pauli. Já a correlação eletrônica está relacionada à interação entre elétrons, em particular a parte dessa interação que é perdida quando se considera que cada elétron interage apenas com o campo médio gerado pelos outros. A correlação eletrônica se deve ao fato de que a probabilidade de se encontrar um elétron nas proximidades de outro é reduzida.

O método de Hartree-Fock ainda é usado como referência, e também serviu de ponto de partida para o desenvolvimento de métodos *ab initio* mais precisos, utilizando diferentes estratégias para uma melhor descrição dos efeitos de correlação eletrônica. Alguns dos métodos pós-Hartree-Fock mais usados são a teoria da perturbação de Moller-Plesset (MP2/MP4), *Configuration Interaction* (CI), e *Coupled Cluster* (CC) [63].

2.2.2 Teoria do Funcional da Densidade

Na teoria quântica, uma vez definido o potencial externo U_{ext} aplicado a um sistema, o problema está inteiramente definido e a função de onda do estado de menor energia $\Psi_0(\mathbf{r})$

pode ser encontrada de maneira unívoca através da equação de Schrodinger (2.1), bem como sua energia E_0 . No caso dos sistemas materiais considerados aqui, isso significa dizer que, para um potencial produzido pelos núcleos atômicos (mais as contribuições externas, se houver), há uma única função $\Psi_0(\mathbf{r})$ que descreve o estado de menor energia dos elétrons, e uma única energia E_0 correspondente a esse estado. Uma vez conhecido esse estado $\Psi_0(\mathbf{r})$, a sua função distribuição de densidade eletrônica $\rho(\mathbf{r})$ pode ser calculada através de $\rho(\mathbf{r}) \equiv \Psi(\mathbf{r})\Psi^*(\mathbf{r})$, onde $\Psi^*(\mathbf{r})$ é o complexo conjugado da função de onda. Logo, na teoria quântica padrão, a ordem natural de cálculo é partir do potencial U_{ext} para calcular a função de onda $\Psi_0(\mathbf{r})$ e sua energia E_0 , e a partir destas, deduzir todas as outras propriedades do sistema, como sua função distribuição de densidade eletrônica $\rho(\mathbf{r})$. Essa relação está ilustrada abaixo:

$$U_{ext} \rightarrow \Psi_0(\mathbf{r}), E_0 \rightarrow \rho(\mathbf{r})$$

A Teoria do Funcional da Densidade (ou *Density Functional Theory*, DFT) se baseia no fato de que, uma vez definida a distribuição de densidade eletrônica $\rho(\mathbf{r})$ do estado de menor energia do sistema, o potencial externo U_{ext} também está univocamente definido, fechando o ciclo da expressão acima [62, 63]. Logo, a função $\rho(\mathbf{r})$ pode ser tratada como a variável fundamental do problema, no lugar da função de onda $\Psi_0(\mathbf{r})$. Em particular, a energia do estado fundamental dos elétrons de um sistema E_0 pode ser determinada completamente a partir de $\rho(\mathbf{r})$. Em linguagem matemática, isso significa que a energia é um funcional da densidade de carga: $E_0 = F[\rho(\mathbf{r})]$. (Um funcional é semelhante a uma função, mas tendo como entrada não um valor numérico, mas uma função.)

2.2.2.1 Teoremas de Hohenberg-Kohn

A DFT se baseia em dois teoremas fundamentais publicados por Hohenberg e Kohn em 1964 [67]. O primeiro teorema afirma que a energia do estado fundamental eletrônico de um sistema é um funcional único da distribuição da densidade eletrônica $\rho(\mathbf{r})$. Ou seja, pode-se escrever $E_0 = F[\rho(\mathbf{r})]$, de maneira que, conhecendo-se a densidade eletrônica $\rho(\mathbf{r})$, pode-se usar o funcional F para se calcular a energia E_0 sem nenhum conhecimento sobre a função da onda do sistema. O segundo teorema afirma que esse funcional F fornece a menor energia possível (a verdadeira energia do estado fundamental do sistema, E_0) se e só se a densidade $\rho(\mathbf{r})$ dada for a verdadeira densidade eletrônica do estado fundamental. Logo, a minimização da energia

através desse funcional (via métodos como o dos multiplicadores de Lagrange) sempre fornece a densidade eletrônica do estado fundamental.

O atrativo da DFT está na simplificação que essa teoria possibilita para problemas quânticos [62, 63]. Encontrar a função de onda de um sistema com N elétrons é um problema de $3N$ dimensões, enquanto que descobrir sua densidade eletrônica total envolve apenas 3. Em teoria, isso permitiria o estudo preciso de sistemas atômicos complexos a um custo computacional muito mais baixo. Por outro lado, o potencial da DFT ainda não pode ser explorado plenamente. Apesar de os teoremas de Hohenberg-Kohn provarem a existência de um funcional $F[\rho(\mathbf{r})]$ que fornece a energia exata de qualquer sistema, eles não dão qualquer pista de como descobrir esse funcional, e até o presente momento sua forma exata segue desconhecida. Dessa forma, a aplicação da DFT depende, atualmente, da utilização de funcionais aproximados.

A primeira etapa na elaboração de funcionais aproximados é dividir a energia em diferentes componentes, como mostrado na Equação 2.5. O termo K corresponde à parcela da energia cinética dos elétrons; U_{ext} representa a interação dos elétrons com o campo externo, incluindo os núcleos atômicos; $U_{classic}$ representa a interação eletrostática clássica entre os elétrons; e E_{XC} inclui os efeitos de troca e correlação eletrônica.

$$E = K + U_{ext} + U_{classic} + E_{XC} \quad (2.5)$$

Dois desses termos, U_{ext} e $U_{classic}$, são de simples tratamento. Eles podem ser facilmente escritos em termos da densidade eletrônica $\rho(\mathbf{r})$, como mostrado nas Equações 2.6 e 2.7. A expressão para U_{ext} depende da posição \mathbf{R} dos núcleos atômicos I , bem como do potencial $v(\mathbf{r} - \mathbf{R}_I)$ entre os núcleos e a densidade eletrônica.

$$U_{ext} = \sum_{I=1}^P \int \rho(\mathbf{r}) v(\mathbf{r} - \mathbf{R}_I) d\mathbf{r} \quad (2.6)$$

$$U_{classic} = \frac{1}{2} \iint \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}' \quad (2.7)$$

Os outros dois termos, entretanto, apresentam dificuldades. O termo cinético é simplificado através das equações de Kohn-Sham, discutidas a seguir, e o termo de troca e correlação é desconhecido, exigindo aproximações (discutidas na Seção 2.2.2.3).

2.2.2.2 Equações de Kohn-Sham

A utilização prática da DFT só se tornou viável após Kohn e Sham sugerirem em 1965 [68] a substituição da energia cinética do sistema multi-eletrônico por aquela do sistema não-interativo equivalente, uma vez que este pode ser facilmente calculado com um campo efetivo, de maneira semelhante ao método de Hartree-Fock. Essa alteração faz com que o termo de energia cinética ignore os efeitos cinéticos de correlação eletrônica, que devem, então, ser considerados em E_{XC} . As funções de onda calculadas para os elétrons não-interativos definem os chamados orbitais de Kohn-Sham ϕ_i , que, apesar de não terem nenhuma proximidade teórica com os auto-estados do sistema eletrônico real, servem como uma boa aproximação em diversos casos [69–72]. Esses orbitais ϕ_i são obtidos a partir das equações de Kohn-Sham (Equação 2.8), que também fornecem as suas energias correspondentes ϵ_i .

$$\left(-\frac{\hbar}{2m}\nabla^2 + v_{ef}(\mathbf{r}) \right) \phi_i = \epsilon_i \phi_i \quad (2.8)$$

O potencial efetivo v_{ef} é dado pela Equação 2.9, que depende do potencial externo v_{ext} , bem como da densidade eletrônica $\rho(\mathbf{r})$ e da derivada do termo de troca e correlação E_{XC} .

$$v_{ef}(\mathbf{r}) = v_{ext}(\mathbf{r}) + e^2 \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \frac{\delta E_{XC}[\rho(\mathbf{r})]}{\delta \rho(\mathbf{r})} \quad (2.9)$$

Utilizando-se os orbitais de Kohn-Sham, obtém-se uma expressão para a energia como funcional da densidade eletrônica, como mostrado na Equação 2.10, estando o potencial externo U_{ext} incluído na modelagem das equações de Kohn-Sham.

$$E_{KS}[\rho(\mathbf{r})] = \sum_{i=1}^N \epsilon_i - \frac{1}{2} \iint \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}' + \left[E_{XC}[\rho(\mathbf{r})] + \int \rho(\mathbf{r}) \frac{\delta E_{XC}[\rho(\mathbf{r})]}{\delta \rho(\mathbf{r})} d\mathbf{r} \right] \quad (2.10)$$

A utilização das equações de Kohn-Sham reintroduz N equações e $3N$ variáveis espaciais ao problema, uma vez que modela orbitais para elétrons individuais. Como resultado, o método apresenta um custo computacional semelhante ao de Hartree-Fock, mas em contrapartida pode fornecer resultados muito mais precisos. De fato, a Equação 2.10 representa uma teoria exata. Todas as aproximações serão feitas no termo de troca e correlação, que aparece entre colchetes na Equação 2.10, e que até o momento não pode ser modelado de maneira exata para o caso geral de problemas eletrônicos.

A seguir são discutidos brevemente os principais tipos de funcionais de troca e correlação aproximados que são usados tradicionalmente. Devido ao fato de que eles utilizam dados empíricos e/ou ajuste a comportamentos-limite ideais em sua formulação, cálculos de DFT são muitas vezes tratados como métodos não puramente *ab initio*, em oposição aos métodos inteiramente baseados em primeiros princípios, como aqueles derivados de Hartree-Fock. Ainda assim, atualmente cálculos DFT são amplamente utilizados em química teórica e ciência computacional de materiais, tendo permitido grandes avanços no estudo de sistemas em escala atômica [61, 63, 73].

2.2.2.3 Funcionais de Troca e Correlação

A aproximação mais simples para troca e correlação é a Aproximação da Densidade Local (*Local Density Approximation*, ou LDA). Essa abordagem considera que a densidade eletrônica varia lentamente, podendo ser considerada localmente homogênea, o que permite tratá-la como um gás homogêneo de elétrons, semelhantemente à teoria de Thomas-Fermi de 1927. A LDA faz boas previsões para sistemas com maior homogeneidade, como sistemas sólidos periódicos [61, 63].

Melhorias à LDA devem considerar um gás não-uniforme de elétrons. Uma maneira de se fazer isso é escrevendo a energia de troca e correlação como uma expansão envolvendo o gradiente da densidade eletrônica, e então truncá-la em certo termo. A LDA mantém apenas o termo de ordem zero (ρ). A inclusão do termo de primeira ordem ($\nabla\rho$) obriga o funcional a seguir certos comportamentos que o funcional exato deveria apresentar, o que leva a aproximações aceitáveis, em muitos casos melhores do que a LDA (mas não para todas as propriedades de todos os sistemas materiais). Essas expansões são chamadas de Aproximações Generalizadas do Gradiente (ou *Generalized Gradient Approximations*, GGA), e estão entre as mais utilizadas atualmente em cálculos de DFT. Alguns exemplos de funcionais de tipo GGA são PW86, LYP e PBE [74].

Uma vez que tanto LDA quanto GGA possuem limitações, outras estratégias foram desenvolvidas para a elaboração de funcionais que levem a uma melhor descrição do sistema, e consequentemente a melhores previsões de propriedades. Uma delas é a adição de mais termos na expansão da densidade eletrônica, resultando nos chamados funcionais meta-GGA. Outra estratégia leva aos chamados funcionais híbridos, que combinam os funcionais discutidos acima com a parcela de troca de acordo com a teoria de Hartree-Fock [61, 63]; alguns exemplos são os

funcionais B3LYP, PBE0 e HSE.

Além disso, há dois tipos principais de interação que cálculos de DFT têm dificuldade em modelar corretamente. O primeiro é a correlação eletrônica forte, presente em materiais altamente correlacionados (como isolantes de Mott, que deveriam ser condutores segundo a teoria convencional de bandas, mas apresentam comportamento isolante). A solução típica para esse problema é o uso da correção semiempírica de Hubbard, denominada DFT+U, que pode ser utilizada com diferentes funcionais de troca e correlação (LDA, GGA). Esse modelo envolve a adição de um termo U ao Hamiltoniano, modelado para capturar a repulsão Coulombiana entre elétrons do mesmo orbital atômico [61, 63]. O custo computacional de DFT+U é apenas ligeiramente maior que o de DFT simples, e permite uma melhor descrição das interações de elétrons d e f , importantes em diversos materiais, como metais de transição.

O segundo problema é a modelagem de interações de van der Waals. Isso se deve à natureza local (LDA) ou semi-local (GGA) dos funcionais comumente utilizados, o que causa uma subestimação das interações de longo alcance, isto é, da dispersão. Uma solução simples e amplamente utilizada é a correção empírica de Grimme, conhecida como DFT-D (D2, D3, etc, dependendo da implementação específica) [75].

2.2.2.4 Outros aspectos de cálculos DFT

Do ponto de vista prático, cálculos de primeiros princípios descrevem orbitais moleculares a partir de expansões de funções ortogonais, empregando variadas bases. A complexidade e/ou o aumento do número de funções de base utilizadas leva a uma melhoria nos resultados, mas a um custo computacional maior. Cálculos em sistemas não-periódicos costumam utilizar funções Gaussianas como base, uma vez que elas são naturalmente localizadas (aproximam-se rapidamente de zero a grandes distâncias). Cálculos em sistemas periódicos, por sua vez, tipicamente utilizam expansões em séries de Fourier empregando ondas planas (combinação linear de senos e cossenos).

Outro fator importante em cálculos DFT é o uso de potenciais de núcleo efetivo, ou pseudopotenciais. As aproximações descritas na Seção 2.2.1 incluíram a modelagem dos núcleos atômicos como partículas clássicas com posições fixas, cujas propriedades são fornecidas como um dado de entrada na simulação, permitindo o cálculo do potencial "externo" que os elétrons experienciam. Na prática, entretanto, essa modelagem faz uso de pseudopotenciais, que representam não apenas o núcleo mas também os elétrons mais internos do átomo, cuja

função de onda (ou densidade eletrônica) é praticamente independente da vizinhança. Dessa forma, apenas os elétrons de valência precisam ser tratados explicitamente durante os cálculos, diminuindo o custo computacional sem que haja grande perda de precisão nos resultados. Essa simplificação não introduz grandes erros ao cálculo, uma vez que os elétrons mais externos são os mais relevantes do ponto de vista prático, isto é, são estes que participam de ligações químicas e definem propriedades eletrônicas, ópticas e magnéticas do material. Um segundo motivo para a utilização de pseudopotenciais é a dificuldade que existe em modelar orbitais internos (principalmente em uma expansão de ondas planas), uma vez que eles são altamente localizados [61, 63]. Existem diferentes classes de pseudopotenciais disponíveis na literatura, cada uma desenvolvida a partir de estratégias específicas; alguns exemplos são *Norm-Conserving*, *Ultrasoft*, e *Projector Augmented Wave* (PAW).

2.2.3 Dinâmica Molecular Clássica

Outro método de simulação atômica amplamente utilizado é a dinâmica molecular (*Molecular Dynamics*, MD) clássica. Em vez de modelar explicitamente os elétrons a partir de equações da mecânica quântica, este método trata cada átomo como uma partícula clássica com posição \mathbf{r} e velocidade \mathbf{v} bem definidas, empregando campos de força (potenciais interatômicos empíricos) para calcular as forças entre eles. Esse cálculo é feito com base na Segunda Lei de Newton, mostrada na Equação 2.11, onde \mathbf{F} é a força total aplicada a cada átomo, \mathbf{r} e m são sua posição e massa, respectivamente, t é o tempo, e U é sua energia potencial (definida pelo campo de força).

$$\mathbf{F} = m \frac{d^2 \mathbf{r}}{dt^2} = -\nabla U \quad (2.11)$$

Mais especificamente, é utilizado um procedimento de integração numérica para simular a evolução temporal $\mathbf{r}(t)$ e $\mathbf{v}(t)$ dos átomos do sistema [76, 77]. Um exemplo comumente empregado para esse fim é o método de Verlet, apresentado nas Equações 2.12 e 2.13, onde Δt é o tempo do passo de integração e \mathbf{a} ($= \frac{\mathbf{F}}{m}$) é a aceleração do átomo.

$$\mathbf{x}(t + \Delta t) \approx 2\mathbf{x}(t) - \mathbf{x}(t - \Delta t) + \mathbf{a}(t)\Delta t^2 \quad (2.12)$$

$$\mathbf{v}(t + \Delta t) \approx \frac{\mathbf{x}(t + \Delta t) - \mathbf{x}(t)}{\Delta t} \quad (2.13)$$

Os campos de força clássicos são parametrizados a partir de dados empíricos e/ou cálculos de primeiros princípios. As forças intermoleculares podem ser modeladas a partir de interações entre pares de átomos, utilizando-se equações analíticas como o tradicional potencial de Lennard-Jones, que representa as forças atrativas de van der Waals e a repulsão de Pauli. Essa expressão está mostrada na Equação 2.14, onde E_{LJ} é a energia potencial da interação, r é a distância entre os dois átomos, e σ e ϵ são parâmetros específicos das espécies em questão (seu elemento e outras características, tais como hibridização), representando, respectivamente, o comprimento e a energia característicos da interação. As forças eletrostáticas E_e , por sua vez, podem ser modeladas através de um potencial de Coulomb que inclua efeitos de blindagem [78].

$$E_{LJ} = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right] \quad (2.14)$$

A modelagem das forças intramoleculares (decorrentes de ligações químicas) também é feita a partir de equações aproximadas, tendo como finalidade capturar a geometria e a rigidez dos modos de vibração das moléculas. Uma vez que dependem da posição de mais de dois átomos, esses campos de força são chamados de potenciais de muitos corpos. Um exemplo típico é utilizar uma aproximação harmônica para modelar ligações químicas covalentes, como mostrado na Equação 2.15, onde E_b é a energia potencial de ligação química, k_r um parâmetro para sua rigidez, r a distância entre os átomos, e r_0 a distância de menor energia. Uma expressão similar (Equação 2.16) pode ser utilizada para modelar a deformação angular, a partir do ângulo θ formado por três átomos consecutivos. Expressões periódicas (Equação 2.17) são usadas para modelar diedros formados por grupos de quatro átomos, capturando a diferença de energia entre conformações alternadas e eclipsadas a partir do ângulo ϕ entre os átomos.

$$E_b = k_r (r - r_0)^2 \quad (2.15)$$

$$E_a = k_a (\theta - \theta_0)^2 \quad (2.16)$$

$$E_d = k_d (1 + \cos(n\phi - d)) \quad (2.17)$$

Cada campo de força usado em MD consiste, portanto, em um conjunto de equações e parâmetros otimizados para um fim específico, como uma classe de sistemas materiais ou

a previsão de uma propriedade em particular. Logo, eles devem ser utilizados com cautela, garantindo-se que a parametrização escolhida seja capaz de modelar os fenômenos de interesse do sistema simulado. Alguns exemplos de potenciais de muitos corpos usados na descrição de sistemas moleculares são CHARMM [79], AMBER [80], COMPASS [81], e OPLS [82]. Também há potenciais de muitos corpos específicos para a simulação de sistemas sólidos, como materiais *bulk* e nanoestruturados, como metais, óxidos, alótropos de carbono, entre outros. Alguns exemplos são ReaxFF [78], AIREBO [83], Tersoff [84], COMB [85], e EAM [86].

Uma vez que a MD clássica ignora a existência dos elétrons, ela é incapaz de prever propriedades eletrônicas. No entanto, esse método permite simulações por intervalos de tempo mais longos e sistemas com maior número de átomos. Por exemplo, simulações MD baseadas em primeiros princípios (*ab-initio*-MD) estão limitadas a picosegundos e poucas centenas de átomos, enquanto a MD clássica chega facilmente a nanosegundos e centenas de milhares de átomos, podendo, em casos extremos, simular microsegundos ou milhões de átomos com auxílio de computação paralela massiva. Portanto, a MD clássica é utilizada para o estudos em maior escala, como por exemplo dinâmica estrutural, termodinâmica de sistemas complexos, fenômenos de transporte, entre outros.

A MD clássica também pode ser utilizada para o estudo de reações químicas, uma vez que foram desenvolvidos potenciais empíricos que computam a criação e a destruição de ligações químicas durante a simulação. Esses campos de força são chamados de potenciais reativos, e alguns exemplos são o ReaxFF (discutido a seguir) e o AIREBO [83]. Assim como os outros potenciais de muitos corpos, os campos de força reativos só devem ser utilizados para simular os tipos de reações químicas que eles foram parametrizados para descrever.

2.2.3.1 ReaxFF

ReaxFF (*Reactive Force Field*) é um dos campos de força reativos mais utilizados na simulação de materiais [87]. Ele combina a modelagem de reações químicas a um baixo custo computacional (comparado a cálculos DFT), o que permite a simulação de sistemas reacionais complexos, como aqueles que ocorrem em interfaces e/ou dependem de fenômenos de transporte.

O ReaxFF utiliza um formalismo de ordem de ligação para modelar a quebra e a formação de ligações entre átomos, simulando reações químicas sem a necessidade de um tratamento eletrônico explícito [78, 88]. A ordem de ligação é computada a partir das distâncias interatômicas, capturando as interações covalentes de longa distância características de estados

de transição, o que permite a predição de energias de ativação reacionais. Além disso, é utilizado um modelo polarizável para as cargas atômicas, atualizando-as a cada passo de tempo. Essas implementações permitem a descrição de interações reativas e não-reativas, tanto de natureza covalente quanto eletrostática, garantindo uma representação suficientemente precisa de uma vasta gama de materiais.

A ordem de ligação BO_{ij} (de *Bond Order*) entre os átomos i e j é calculada através da Equação 2.18, onde r_{ij} é a distância interatômica, r_0 são as distâncias de equilíbrio das ligações, e p_i são parâmetros empíricos. As três contribuições são relativas ao caráter σ , π e $\pi\pi$ da ligação. A equação é contínua, o que permite transições suaves entre as ordens de ligação.

$$BO_{ij} = BO_{ij}^{\sigma} + BO_{ij}^{\pi} + BO_{ij}^{\pi\pi} = \exp \left[p_1 \left(\frac{r_{ij}}{r_0^{\sigma}} \right)^{p_2} \right] + \exp \left[p_3 \left(\frac{r_{ij}}{r_0^{\pi}} \right)^{p_4} \right] + \exp \left[p_5 \left(\frac{r_{ij}}{r_0^{\pi\pi}} \right)^{p_6} \right] \quad (2.18)$$

Como a maioria dos potenciais de muitos corpos, o ReaxFF possui termos de energia para ligações (E_b), ângulos (E_a) e diedros (E_d), além de interações de van der Waals (E_{vdW}) e de Coulomb (E_e), todas descritas anteriormente. Em particular, o cálculo das forças de van der Waals é feita através do potencial de Morse [78]. Todos esses termos estão listados na Equação 2.19, que modela a energia total de um sistema. Além desses termos, o ReaxFF possui uma penalidade de energia E_{over} , que impede a super-coordenação dos átomos com base em suas regras de valência. O termo $E_{specific}$ representa interações específicas do sistema, e em geral não precisa ser incluído.

$$E_{ReaxFF} = E_b + E_a + E_d + E_{vdW} + E_e + E_{over} + E_{specific} \quad (2.19)$$

A parametrização do ReaxFF é feita a partir de cálculos eletrônicos, em geral de DFT, o que permite que novos sistemas sejam parametrizados de forma sistemática. Atualmente, diversas parametrizações estão disponíveis para o ReaxFF. Elas podem ser divididas em dois grupos principais, o de combustão e o de sistemas aquosos, além de parametrizações independentes [87].

As diversas parametrizações do ReaxFF permitem a simulação de uma vasta gama de fenômenos reacionais, como catálise heterogênea, deposição atômica, fenômenos eletroquímicos, nanoindentação, degradação térmica, oxidação de superfícies, adsorção, conformação de biomoléculas, entre outros. Além disso, o formalismo da quebra de ligações torna possível a

modelagem da fratura de materiais, o que possibilita a simulação de ensaios para a investigação de propriedades mecânicas.

2.3 ALGORITMOS GENÉTICOS

Métodos de otimização numérica estão no coração da ciência e da engenharia. No caso mais simples, tem-se uma função objetivo $y = f(x)$ e deseja-se saber qual é o valor da entrada x que maximiza (ou minimiza) a saída y . Em casos como esse, em geral é possível encontrar a solução de maneira analítica. Mais comumente, entretanto, a grandeza a ser otimizada depende de diversas variáveis, podendo apresentar diferentes máximos locais, além de outras irregularidades e limites de aceitabilidade. Nesses casos, quase sempre é necessário um algoritmo de otimização numérica, dos quais alguns exemplos clássicos são o método de Newton e variantes, como o método do gradiente conjugado.

Em problemas ainda mais complexos, o número de máximos locais é tão grande (e difícil de estimar) que as soluções fornecidas por algoritmos simples não são confiáveis. Nesses casos é inviável realizar uma exploração completa (ou mesmo extensiva) da função objetivo, sendo necessário um algoritmo capaz de encontrar máximos globais a partir de uma amostragem relativamente pequena de pontos. Um exemplo dessa classe de problemas é o design de materiais: existe um número praticamente infinito de estruturas atômicas metaestáveis, produzidas com diferentes geometrias e espécies atômicas – ou, vendo-se pelo lado experimental, inúmeras combinações de parâmetros de síntese que levam a morfologias diferentes. Há diversos métodos de simulação que podem auxiliar no design de materiais, como Monte Carlo, metadinâmica, recozimento simulado (*simulated annealing*), saltos de mínimos (*minima hopping*), saltos de bases (*basin hopping*), entre outros [89, 90].

Uma classe de métodos capazes de resolver problemas complexos de otimização são os chamados algoritmos evolutivos, que utilizam os princípios da evolução biológica para favorecer entradas (x) que produzem valores altos de saída (y). Algumas das classes de algoritmos evolutivos mais utilizadas em ciência e engenharia de materiais são os algoritmos genéticos, introduzidos por Holland [91] e Rechenberg [92], e os de otimização por enxame de partículas (*particle swarm optimization*). Um exemplo deste último é o software CALYPSO, que é distribuído livremente para uso acadêmico e já levou à descoberta de diversos novos materiais [93].

Algoritmos genéticos associam cada material (fenótipo) a um conjunto de dados (genótipo ou genoma) que codifica suas características relevantes, tais como as posições atômicas em uma célula unitária ou os parâmetros utilizados na sua síntese. É gerada uma população de materiais individuais (candidatos), para os quais é medido, experimental ou computacionalmente, o valor de alguma propriedade de interesse (aptidão ou *fitness*) – por exemplo, um parâmetro de estabilidade relativa. Em seguida, os candidatos mais aptos são selecionados para participar de cruzamentos e mutações, nos quais seus códigos genéticos são cruzados ou levemente alterados, respectivamente, produzindo novos candidatos para a próxima geração (Figura 5). Depois de algumas gerações, o algoritmo terá favorecido os genes que maximizam a propriedade de interesse, com grande chance de encontrar não apenas o máximo global, mas também uma variedade de máximos locais, que podem representar materiais metaestáveis de interesse científico e tecnológico.

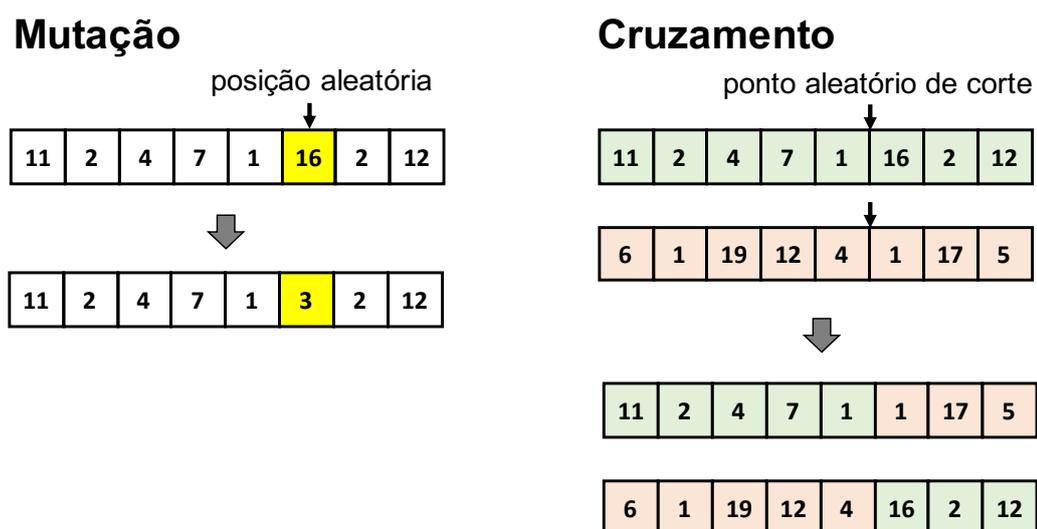


FIGURA 5 – Operações de mutação e cruzamento realizadas em genótipos ilustrativos. O genótipo é um conjunto de números que codificam as características dos candidatos. A mutação altera um único gene, enquanto o cruzamento corta o código genético de dois candidatos e os cruza para produzir dois novos candidatos. Uma explicação mais detalhada dessas operações é dada na Seção 2.3.1.3.

A utilização de algoritmos genéticos no design experimental de novos materiais teve início com o trabalho pioneiro de Wolf, que aplicou a estratégia evolutiva à descoberta de catalisadores heterogêneos [94]. A aptidão de cada material foi definida como a sua produção de propeno, que foi medida experimentalmente. O genótipo foi usado para codificar a composição de cada catalisador. O artigo reportou resultados promissores, demonstrando que algoritmos evolutivos podem ser mais eficientes no design de materiais comparado a outros métodos estocásticos. A aplicação de algoritmos genéticos para o design computacional de materiais,

por sua vez, começou a produzir resultados promissores com o trabalho de Deaven e Ho, que otimizaram a geometria de clusters de carbono, produzindo fulerenos de até 60 átomos com uma performance superior à do método de recozimento simulado [95].

Ainda hoje, a descoberta e otimização de catalisadores é a principal aplicação dos algoritmos genéticos à engenharia de materiais [89]. Alguns outros exemplos são a descoberta de novas estruturas cristalinas [96], a otimização de clusters atômicos [97], o estudo da distribuição de defeitos pontuais [98], entre outros [89]. Nesses casos são utilizados tanto métodos experimentais quanto computacionais para a medição da aptidão dos candidatos. A próxima seção trata em maior detalhe do segundo caso, no qual simulação computacional é utilizada na modelagem dos materiais de interesse.

2.3.1 Algoritmos genéticos aplicados ao design computacional de materiais

O design computacional de materiais envolve a utilização de métodos atomísticos, como simulações de MD clássica e cálculos DFT, na modelagem de estruturas atômicas. No contexto de algoritmos evolutivos, uma abordagem puramente computacional significa utilizar essas ferramentas para o cálculo da aptidão dos candidatos – isto é, alguma propriedade de interesse.

No caso geral de algoritmos evolutivos, o espaço de possíveis candidatos é chamado de paisagem de aptidão (*fitness landscape*), e a aptidão é tida como uma variável a ser maximizada. No design computacional de materiais, entretanto, muitas vezes fala-se de uma superfície de energia potencial (*potential energy surface*, PES) ou paisagem de energia (*energy landscape*), na qual a energia deve ser minimizada para se encontrar os candidatos mais estáveis [99, 100]. Uma vez que os problemas de maximização e minimização são computacionalmente equivalentes, o mesmo tratamento é dado a ambos os casos.

Trabalhos pioneiros demonstraram o potencial dos algoritmos genéticos no design computacional de materiais, com performances superiores às de outros métodos estocásticos [95, 101]. Uma vez que os métodos computacionais utilizados permitem uma descrição a nível atômico dos materiais de interesse, as principais aplicações dessa metodologia são na otimização de clusters atômicos [97] e na descoberta de novas fases cristalinas [96]. Em alguns casos, é possível a previsão de todo o diagrama de fases de um sistema, com variação na composição dos candidatos e cálculo da energia relativa de formação através do método de casco convexo (*convex hull*) [102].

Diversos algoritmos genéticos desenvolvidos nas últimas duas décadas possuem interface com pacotes de MD e/ou DFT, o que permite sua utilização direta em ambiente computacional. A grande maioria deles tem como finalidade a previsão de novas estruturas cristalinas. Em 2006, Oganov e Glass reportaram o algoritmo USPEX (*Universal Structure Predictor: Evolutionary Xtallography*), capaz de prever estruturas cristalinas estáveis a partir somente de sua composição [100]. Em 2013, Brahmman e Kortus publicaram o programa EVO, que incluía a predição de estruturas 2D [103]. Ainda em 2013, Tipton e Henning reportaram o GASP (*Genetic Algorithm for Structure and phase Prediction*), primeiro a permitir a variação na composição dos candidatos [102]; e que em 2016 foi expandido para o estudo de estruturas 2D [104]. Em 2015, Mrozek et al. publicaram o HPEA (*Hybrid Parallel Evolutionary Algorithm*) [105]. Em 2021, Falls et al. fizeram uma revisão sobre métodos de predição a priori de estruturas cristalinas, com enfoque no estado atual de seu próprio algoritmo evolutivo, o XtalOpt [90]. Outros exemplos são os algoritmos de Abraham e Probert [106], de Trimarchi e Zunger (GSGO) [107], de Wu et al. [108], e de Fadda e Fadda [109].

Também foram propostos algoritmos genéticos para a otimização de clusters atômicos [97], sendo um dos mais utilizados o de Vilhelmsen e Hammer [99], que permite a relaxação via DFT de estruturas suportadas em superfícies, incluindo materiais porosos. Esse algoritmo faz parte do Ambiente de Simulação Atômica (*Atomic Simulation Environment, ASE*), um pacote disponível em Python [110]. Outros grupos desenvolveram seus próprios algoritmos genéticos para a exploração de conjuntos mais restritos de materiais e/ou para aplicações específicas. Isso inclui o design de catalisadores sólidos [101] e nanoestruturados [111]. Além disso, foram desenvolvidos algoritmos específicos para a predição de cristais moleculares, nos quais as moléculas têm suas posições alteradas sem se quebrarem [112].

Entre os feitos dos algoritmos genéticos no design computacional de materiais estão a predição de novas fases supercondutoras e semicondutoras [113], bem como de novos dióxidos 2D do grupo IV para aplicações dielétricas [114] e folhas de boro [115]. Foram estudados diversos sistemas, como Si, SiC, GaAs e AuPd [107]. Foi feita a previsão de diversas nanopartículas *core-shell* para catálise da reação de redução do oxigênio (*Oxygen Reduction Reaction, ORR*) [111]. Também foram desenvolvidos óxidos metálicos em diferentes composições para a desidrogenação catalítica do propano [101].

As próximas Subseções tratam em maiores detalhes as diversas partes de um algoritmo genético, discutindo as estratégias já desenvolvidas para utilizá-lo no design computacional de

materiais.

2.3.1.1 Genótipos

A codificação de cada material em um genótipo deve ser feita de maneira que uma alteração nos genes (a entrada x) produza uma variação correspondente no valor de aptidão (a saída y). Idealmente, a variação na aptidão deve ser mais ou menos proporcional à alteração produzida no genótipo: pequenas mutações devem causar pequenas mudanças na aptidão, e em cruzamentos os candidatos filhos devem ter aptidões parecidas com a de seus pais (devem "ser parecidos" com eles) [95]. Para que isso ocorra, os parâmetros codificados devem ser aqueles que de fato regem a propriedade de interesse, preferencialmente da maneira mais desacoplada possível (com cada pedaço do genótipo afetando o valor de aptidão de maneira independente).

Em algoritmos genéticos de aplicação geral, o genótipo costuma ser codificado como uma série binária, que pode ser cortada em qualquer ponto durante um cruzamento [89]. Isso garante a universalidade do algoritmo, que pode, então, ser utilizado para qualquer tipo de candidato. Em ciência computacional de materiais, entretanto, a representação de um sistema costuma se dar através de uma lista de espécies e posições atômicas. Logo, essa é uma escolha natural para o genótipo de um candidato. Além disso, trabalhos pioneiros demonstraram que uma representação em espaço real melhora a performance do algoritmo [95, 106]. Isso significa tratar cada posição atômica como um gene, ou seja, uma informação irreduzível, impedindo que uma coordenada seja "cortada ao meio" durante um cruzamento, o que poderia produzir um filho com características radicalmente diferentes de ambos os pais (Figura 6). Em um tratamento como esse (conservando informações irreduzíveis), a representação binária se torna não apenas desnecessária, mas prejudicial, uma vez que mascara o significado de cada gene. Em suma, o algoritmo se beneficia de representações genéticas que reflitam da maneira mais orgânica possível o sistema representado.

Diversos algoritmos genéticos utilizam a metodologia descrita acima, permitindo a variação das posições atômicas de maneira a explorar candidatos sem um viés estrutural. Essa estratégia foi utilizada inicialmente na otimização de clusters atômicos [95, 97, 99, 116], mas seu potencial é mais bem explorado no caso de materiais cristalinos [90, 100, 102–107, 109, 113], uma vez que permite a exploração completa de possíveis estruturas com uma dada composição. Em alguns casos, ainda, o algoritmo permite variações na composição do sistema, sendo possível prever todo o diagrama de fases de dois ou mais componentes [102, 104]. A representação

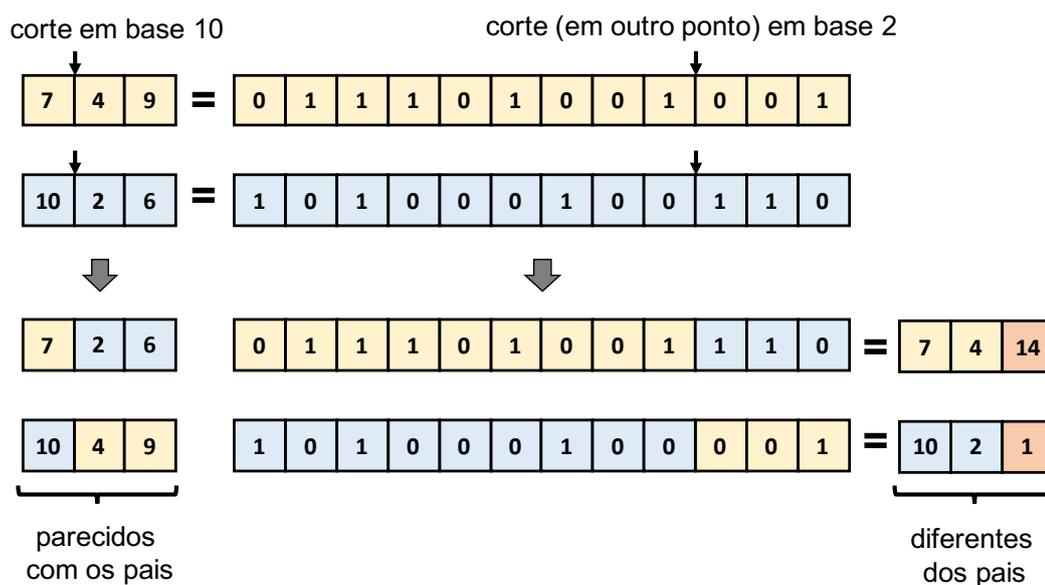


FIGURA 6 – Cruzamento realizado em genótipos codificados em base 10 e base 2.

do sistema em coordenadas reduzidas facilita o cruzamento de materiais com células unitárias diferentes, permitindo uma otimização geométrica juntamente à das posições atômicas [106]. Além disso, no estudo de sistemas bidimensionais, pode-se codificar apenas as coordenadas x e y dos átomos [103–105], permitindo-se (ou não) que a posição z varie durante um processo de relaxação estrutural.

Em alguns casos, entretanto, deseja-se fazer uma exploração em menor escala – em geral quando já existe algum conhecimento empírico sobre os materiais de interesse, o que permite uma maior restrição de sua estrutura. Nesses casos, é benéfico utilizar genótipos minimalistas, que codificam apenas as variáveis que se quer explorar. Como um contra-exemplo do que foi discutido acima, o estudo de vacâncias em um plano cristalino tem uma representação naturalmente binária e bidimensional, sendo útil a utilização de uma tabela de uns e zeros como genótipo [98] (Figura 7). Em um trabalho sobre nanopartículas *core-shell* [111], o número e posição dos átomos foram fixados, codificando-se apenas duas variáveis (genes) no genótipo: a espécie atômica interna (*core*) e externa (*shell*). Nesse caso, o cruzamento pode produzir apenas quatro candidatos filhos (Figura 8).

2.3.1.2 Cálculo da Aptidão

A estimativa da aptidão pode ser feita a partir de experimentos ou de simulações computacionais. No primeiro caso, o algoritmo genético é empregado em paralelo com os experimentos, com alternância entre a construção de novas gerações e as medidas para estimar

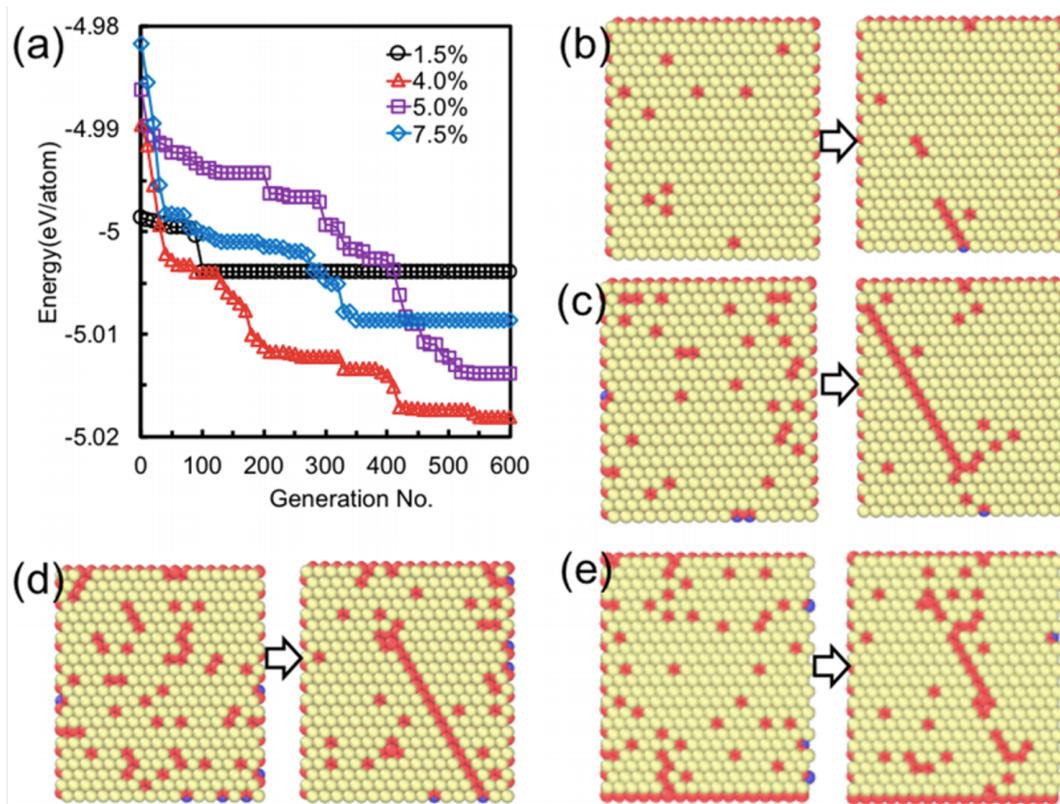


FIGURA 7 – Simulação evolutiva de vacâncias de S em MoS_2 . (a) Energia diminuindo com a passagem de gerações. (b-e) Representação bidimensional do sistema com diferentes números de vacâncias. Reproduzido com permissão de Patra et al. [98], Copyright 2018 American Chemical Society.

a aptidão de cada candidato. Na estimativa computacional, por outro lado, todas as etapas do algoritmo podem ser realizadas de maneira automática, com a implementação do fluxo de trabalho apropriado. Nesses casos a propriedade de interesse é medida através de algum método de modelagem e simulação, como aqueles discutidos na Seção 2.2, como MD clássica reativa [98, 105] e DFT (sendo DFT-GGA o caso mais comum [99, 104, 106, 111, 113], mas havendo também uso de DFT-LDA [107]). Em alguns casos é utilizada uma combinação de métodos clássicos e quânticos para diminuir o tempo de cálculo sem perda de precisão [108]. A aptidão também pode ser estimada com um método de *machine learning*. Redes neurais artificiais, por exemplo, vêm sendo utilizadas com frequência crescente na previsão das propriedades de materiais. Esse método se torna uma ferramenta poderosa quando utilizada para estimar a aptidão de candidatos dentro de uma metodologia evolutiva, uma vez que substitui a necessidade de simulações atômicas, possibilitando a avaliação de um grande número de candidatos em um curto período de tempo – ainda que seja necessário treinar a rede neural antes [101].

A aptidão pode ser qualquer propriedade que se deseje maximizar (ou minimizar) nos

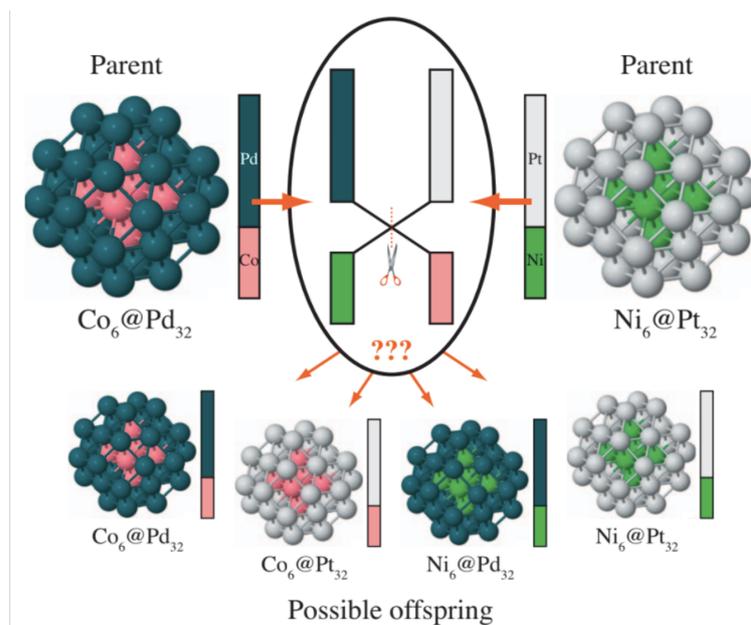


FIGURA 8 – Nanopartículas *core-shell* de dois genes, bem como todos os possíveis cruzamentos entre elas. Reproduzido de Froemming e Henkelman [111], com permissão da AIP Publishing.

materiais em estudo. Uma vez que métodos de simulação computacional permitem o cálculo de diversas propriedades termodinâmicas, a aptidão é geralmente uma medida da energia do sistema, que deve ser minimizada para otimizar sua estabilidade. A energia relativa pode ser calculada com variados termos de contribuição e expressa de diferentes formas, dependendo do método disponível e da precisão desejada. Pode-se utilizar valores de entalpia [106], energia livre de Gibbs [103], energia de formação [104, 113], entre outros. Em outros casos, quando deseja-se otimizar materiais para um fim específico, a aptidão pode ser uma medida de outra propriedade que não a estabilidade. Por exemplo, o design computacional de catalisadores pode ser feito calculando-se sua estrutura de bandas, via DFT, e comparando-se as energias eletrônicas às de um material conhecido com boas propriedades catalíticas [111]. Também é possível utilizar múltiplas propriedades no cálculo da aptidão, caso no qual é preciso ponderá-las arbitrariamente.

Uma vez calculada a propriedade de interesse, pode-se utilizar seu valor diretamente como aptidão ou, mais comumente, normalizá-lo e tratá-lo com alguma equação. A normalização geralmente é feita com algum valor de referência ou utilizando-se o intervalo de aptidões da geração atual, de maneira que os valores tratados de aptidão fiquem entre zero e um [99]. Nesse caso a aptidão efetiva é linearmente proporcional ao valor da propriedade calculada; mas também pode-se utilizar outros tipos de equação, na forma de uma função exponencial ou tangente hiperbólica, o que dá maior ou menor peso às variações da propriedade bruta, representando diferentes graus de pressão evolutiva [99, 116].

Diversas restrições podem ser impostas aos materiais em estudo, caso haja conhecimento prévio que as justifique. Nos casos em que a aptidão é medida experimentalmente e o genótipo codifica as condições de síntese, por exemplo, pode-se restringir a temperatura de síntese a uma faixa de razoabilidade. No caso computacional em que os genes são coordenadas atômicas em uma célula unitária, é possível impedir que dois átomos estejam próximos demais ou que os vetores da célula formem ângulos muito grandes ou pequenos [103, 107].

A maioria dos pacotes de simulação (MD, DFT, etc) também podem fazer otimizações estruturais (relaxações), nas quais as posições atômicas são otimizadas em busca do mínimo local de energia. Essa relaxação pode ser empregada antes do cálculo da aptidão, garantindo que cada candidato corresponda a um material metaestável (com um grau maior ou menor de estabilidade) [103, 105]. Nos casos em que o genoma codifica posições atômicas, essa otimização faz com que o fenótipo deixe de coincidir com o genótipo; nesses casos, pode-se atualizar o genoma com as coordenadas relaxadas, ou não. Caso haja atualização, diz-se que o código genético "sofre alterações durante a vida do candidato", o que corresponde a uma evolução Lamarckista, enquanto o oposto faz uso de conceitos puramente Darwinianos [117]. Essas otimizações locais, apesar de exigirem um tempo computacional maior, podem melhorar a performance do algoritmo genético, uma vez que reduzem consideravelmente a paisagem efetiva de aptidão – isto é, o algoritmo precisa lidar apenas com mínimos locais.

Um algoritmo genético pode ter diversos critérios de parada. O mais comum é um número fixo de gerações exploradas, mas também pode-se identificar quando os valores de aptidão estabilizam (com variações abaixo de um determinado valor), ou quando a aptidão alcança um valor limite [105].

2.3.1.3 Formação de uma nova geração

A última etapa de um algoritmo genético é a construção de uma nova geração a partir dos dados de aptidão da anterior. Há diversas operações envolvidas nesse processo, como seleção de candidatos, cruzamento e mutação. A seleção deve ser feita de maneira a privilegiar os candidatos mais aptos – ou seja, a probabilidade de um material ser selecionado deve ser proporcional à sua aptidão. No entanto, isso pode ser feito através de diferentes métricas e estratégias, sendo possível a inclusão de um termo de singularidade para ajudar o algoritmo a explorar áreas diferentes da paisagem de aptidão. A seguir, os candidatos selecionados passam por processos de cruzamento e/ou mutação antes de seguir para a próxima geração, havendo

também a possibilidade de serem transferidos diretamente, sem alteração genética (um processo chamado de replicação ou elitismo).

Como mencionado na Seção anterior, a aptidão pode ser normalizada e tratada através de diferentes equações, o que define a probabilidade de seleção dos candidatos. Além disso, diferentes estratégias podem ser empregadas na seleção. O método mais comum é o da roleta (*roulette wheel*), no qual os candidatos são selecionados a partir de probabilidades previamente calculadas (proporcionais à aptidão de cada um). Outras metodologias também são possíveis, como o método de torneio, no qual são formados diversos grupos de candidatos dos quais os mais aptos são selecionados [116]. Uma vez selecionado, um candidato pode ser passado para a próxima geração sem alterações genéticas (replicação), ou depois de passar por cruzamentos e/ou mutações.

O processo de replicação pode se dar de maneira estocástica ou determinística. A primeira foi descrita acima, na qual cada candidato tem uma probabilidade de replicação proporcional à sua aptidão. No segundo método, o determinístico, os candidatos mais aptos são sempre replicados na próxima geração, o que garante que cada geração possui um máximo de aptidão igual ou superior à anterior. O método estocástico ajuda algoritmos genéticos a escapar de mínimos locais de energia, sendo ideal para a exploração de polimorfos [106]. Em ambos os casos, pode-se programar o algoritmo genético para não calcular duas vezes a aptidão de um mesmo candidato, o que salva algum tempo computacional. A replicação não fornece um novo material à exploração; sua finalidade é perpetuar os genes mais bem sucedidos, dando-lhes mais chances de participar de mutações e cruzamentos no futuro.

No cruzamento (*crossover*), dois candidatos (pais) produzem um ou dois novos materiais (filhos). Os genótipos dos pais são "cortados" em um ponto aleatório, então cruzados de maneira a produzir dois filhos, como mostrado anteriormente na Figura 5. Dessa forma, dois novos candidatos são gerados. Com sorte, um deles combinará as boas características estruturais de ambos os pais, apresentando uma aptidão superior à de ambos. Já na mutação (*mutation*), um único candidato tem seu genótipo alterado aleatoriamente – em geral de maneira mínima (Figura 5). Enquanto o cruzamento tem como objetivo um grande salto na paisagem de aptidão, evitando assim que o algoritmo fique preso em máximos locais, a mutação causa alterações menores, permitindo que os materiais sejam otimizados localmente.

Tanto no cruzamento quanto na mutação, é preciso que o material filho tenha um

genótipo plausível. Isso significa, por exemplo, garantir que não haja átomos sobrepostos na célula unitária, ou que a temperatura de síntese esteja dentro de um intervalo de aceitabilidade. Uma vez escolhida uma classe de materiais para estudo, bem como a codificação de suas características em um genótipo, é possível modelar as operações de cruzamento e mutação de maneira que elas respeitem as particularidades dessa representação.

Em função disso, tanto a mutação quanto o cruzamento podem ocorrer de diferentes maneiras, dependendo do sistema em estudo. Genótipos que codificam a posição dos átomos devem ser editados em espaço real para que o algoritmo genético tenha algum grau de sucesso [95]; isso significa, por exemplo, utilizar um mesmo plano de corte para cruzar dois candidatos, o que pode ser precedido pela rotação de um deles. Além disso, sistemas periódicos devem sofrer cruzamentos periódicos para garantir que não haja descontinuidades [106]. Como mencionado anteriormente, a representação de coordenadas reduzidas permite o cruzamento de materiais com células unitárias diferentes. O cruzamento também pode ser feito com um ou dois pontos de corte [116], como ilustrado na Figura 9; e com ou sem a restrição de que aproximadamente 50% do genótipo de cada pai seja utilizado, garantindo que nenhum deles seja favorecido [106].

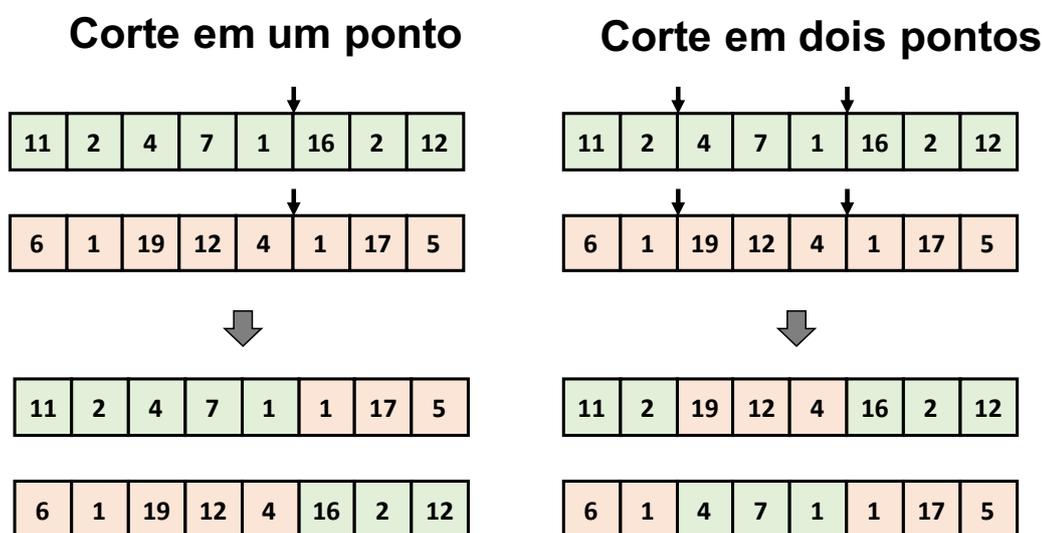


FIGURA 9 – Cruzamento com corte em um e dois pontos.

Diferentes tipos de mutação também são possíveis. Pode-se alterar um único gene ou mais de um, ou trocar dois genes de lugar [111]. Pode ser feita uma translação de parte dos átomos, uma rotação, uma permutação do número atômico de parte dos átomos, ou o espelhamento de parte dos átomos [99]. Além disso, a mutação pode ser feita de maneira independente ou como uma segunda etapa do cruzamento, com o objetivo de introduzir novos genes no processo evolutivo. No último caso, é importante que nem todos os candidatos produzidos por cruzamento

sejam mutados, uma vez que isso pode introduzir muito "ruído" à evolução, o que prejudica a convergência do algoritmo [99].

Há diversas outras variações metodológicas que podem ser levadas em consideração na elaboração de algoritmos genéticos. Uma delas é uma população de tamanho variável, ou seja, gerações com números diferentes de candidatos – ou ainda o abandono da ideia de geração, trabalhando-se em vez disso com um conjunto constantemente variável de candidatos (uma população), o que facilita a paralelização dos cálculos de aptidão [90]. Também pode ser utilizado um modelo de ilhas, no qual diversas subpopulações são simuladas separadamente, permitindo que diferentes características positivas sejam otimizadas em paralelo [118]. Essas subpopulações podem então ser combinadas em uma única população (um continente), onde ocorre a mistura dos genes mais promissores.

Outra medida que garante a diversidade dos candidatos, impedindo que o algoritmo fique preso em um mínimo local, é a implementação de algum fator de singularidade (*uniqueness*), que quantifica o quão diferente cada candidato é dos demais. Essa medida de singularidade pode então ser considerada junto da aptidão no cálculo da probabilidade de seleção dos candidatos, favorecendo materiais com genótipos "novos" [99]. Isso é particularmente importante nos casos em que os genes guardam valores contínuos, como posições atômicas, uma vez que genótipos diferentes podem simbolizar o mesmo material ou materiais muito parecidos, os quais não há interesse em explorar. Além dessas pequenas variações numéricas, as simetrias do sistema impedem que o próprio genótipo seja utilizado na comparação dos candidatos, sendo necessária uma métrica que represente univocamente cada material. Um método simplório é a comparação de energias [116], que pode gerar falsos positivos, uma vez que estruturas muito diferentes podem ter energias próximas [99]. Métodos mais bem sucedidos incluem funções de distribuição radial (*Radial Distribution Function*, RDF) e padrões de difração de raios X, que funcionam como "digitais" dos materiais uma vez que são invariantes com respeito a operações de simetria, como translação e rotação [100].

Outros métodos podem ser usados para garantir a diversidade de uma exploração evolutiva. O programa EVO, por exemplo, traz a noção de "idade" de um candidato: o número de gerações pelas quais ele se replicou [103]. Esse algoritmo restringe a idade dos candidatos a um valor máximo (uma "expectativa de vida"), impedindo que eles se repliquem indefinidamente. A já citada representação padronizada de células unitárias também ajuda a filtrar materiais iguais, uma vez que facilita a comparação de sistemas que têm representações diferentes, mas são

equivalentes por operações de simetria [90].

3 MÉTODOS COMPUTACIONAIS

3.1 ANÁLISE DAS PROPRIEDADES DE NANOFILAMENTOS DE CARBONO DERIVADOS DE MOLÉCULAS HETEROCÍCLICAS DE CINCO MEMBROS

Cálculos baseados na Teoria do Funcional da Densidade (*Density Functional Theory*, DFT) foram realizados para o estudo das propriedades estruturais, eletrônicas e mecânicas de NTHs derivados de compostos aromáticos heterocíclicos de cinco membros – especificamente, moléculas de tiofeno, furano, pirrol e fosfol. As estruturas estudadas estão ilustradas na Figura 10 para o caso do furano. Elas se chamam *syn*, *anti* e *syn-anti*, dependendo do padrão de ligações e da orientação relativa dos grupos anexados (as diferenças são melhor discutidas na Seção 4.1.1). Os NTHs derivados de outras moléculas de cinco membros possuem as mesmas estruturas, substituindo-se apenas o oxigênio pelo grupo funcional correspondente (S, NH e PH nos casos de tiofeno, pirrol e fosfol, respectivamente).

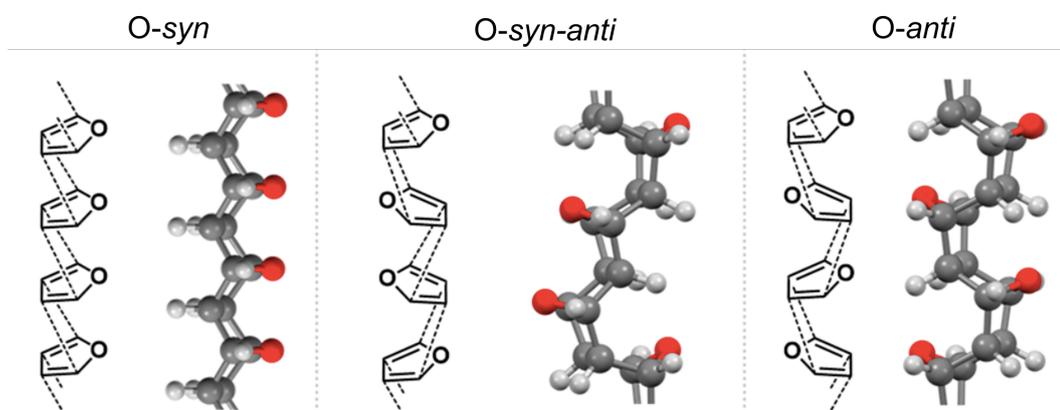


FIGURA 10 – Estruturas de NTHs derivados de furano. Adaptado com permissão de Huss et al. [7], Copyright 2021 American Chemical Society.

Os cálculos de DFT foram realizados no pacote *open-source* Quantum Espresso [119]. Foi utilizado o funcional GGA/PBE para troca e correlação [74], bem como a correção D2 de Grimme para as interações de van der Waals [75]. Foram empregados pseudopotenciais do tipo *norm-conserving* para a representação dos núcleos e elétrons internos, e ondas planas para a expansão dos orbitais de Kohn-Sham (com uma precisão de 80 Ry). A malha de Monkhorst-Pack usada para mapear a primeira zona de Brillouin foi de $1 \times 1 \times 9$ (ou de $1 \times 1 \times 5$, para as estruturas do tipo *syn-anti*, cuja célula unitária é maior). Além disso, garantiu-se um vácuo de 15 Å nas direções perpendiculares ao nanofilamento, com a finalidade de evitar interações entre

a estrutura e suas réplicas no sistema periódico. A convergência foi verificada para todos os parâmetros, que são similares aos utilizados em trabalhos semelhantes da literatura [9, 42].

As posições atômicas de cada estrutura foram relaxadas, bem como a dimensão axial da célula, de maneira a buscar a configuração de menor energia. A seguir, foram calculadas a estrutura eletrônica de bandas, a densidade de estados eletrônicos (*density of states*, DOS) e a densidade de estados eletrônicos projetada em orbitais (*projected DOS*, PDOS) para as estruturas otimizadas. Todos esses cálculos foram feitos no nível DFT GGA/PBE, conhecido por subestimar os valores de *band gap* eletrônico de materiais. Portanto, também foram computados *band gaps* utilizando-se o funcional híbrido de troca e correlação HSE, buscando valores mais precisos. Para isso, utilizou-se uma malha de pontos q com a mesma precisão daquela utilizada em pontos k (isto é, $1 \times 1 \times 5$ ou $1 \times 1 \times 9$, como mencionado acima). Os cálculos realizados com o funcional híbrido HSE partiram das coordenadas atômicas otimizadas com DFT GGA/PBE.

Propriedades mecânicas foram computadas a partir da simulação de testes de deformação uniaxial, utilizando-se uma metodologia tradicional da literatura, que consiste em relaxar as posições atômicas para diferentes valores constantes de deformação aplicada à supercélula [9, 42]. Ao final da otimização das estruturas, calculou-se a tensão na direção correspondente, e levantaram-se curvas de tensão-deformação, a partir das quais foram extraídas algumas propriedades mecânicas. O Módulo de Young foi calculado como a inclinação da curva de tensão-deformação na região linear (até 4% de deformação). A tensão e deformação de ruptura (*ultimate tensile strenght* e *fracture strain*) correspondem ao ponto máximo das curvas de tensão-deformação, que em todos os casos corresponde ao ponto de ruptura (NTHs apresentam comportamento frágil). As propriedades foram expressas em termos de tensão 1D – isto é, unidades de força (nN) – para evitar a necessidade de se definir uma área de seção transversal para os filamentos, requerida para cálculo da tensão 3D convencional (comumente apresentada em GPa). A definição dessa área é uma escolha arbitrária para materiais unidimensionais em geral, devido aos seus diâmetros característicos na mesma ordem de distâncias interatômicas. A tensão de ruptura 1D, por sua vez, corresponde à força necessária para se romper um único nanofilamento.

Também foram realizadas simulações de MD *ab initio* de estruturas selecionadas, com a finalidade de avaliar sua estabilidade térmica. As simulações foram feitas a 300K com um passo de tempo de 1.0 fs durante o tempo necessário para que a energia do sistema se estabilizasse (de 5.0 a 20.0 ps). Foi utilizado o integrador de Verlet para a evolução temporal do sistema, e o

termostato de Berendsen para controle da temperatura. A energia e a tensão do sistema foram monitoradas para garantir que sua estabilidade não era comprometida durante a simulação.

3.2 DESENVOLVIMENTO DE UM ALGORITMO GENÉTICO PARA EXPLORAÇÃO DE ESTRUTURAS COMPLEXAS DE NANOFILAMENTOS DE CARBONO

Com o objetivo de possibilitar o estudo sistemático de estruturas complexas de NTHs através de métodos de simulação computacional, foi desenvolvido um algoritmo genético em Python para a otimização de materiais derivados de moléculas aromáticas. O algoritmo foi implementado a partir do zero porque, apesar de haver pacotes disponíveis para a evolução de sistemas atômicos (Seção 2.3), NTHs possuem restrições estruturais específicas devido à sua síntese a partir de moléculas precursoras (Seção 2.1), sendo necessária uma estratégia específica para sua codificação genética e evolução. A Figura 11 ilustra como NTHs são formados a partir de moléculas de benzeno. As classes II e III dizem respeito à distribuição das ligações de cada molécula, onde a razão do número de ligações feitas para cima e para baixo pode ser 2:4 ou 3:3 (a divisão 1:5 resulta em estruturas de alta energia [4]).

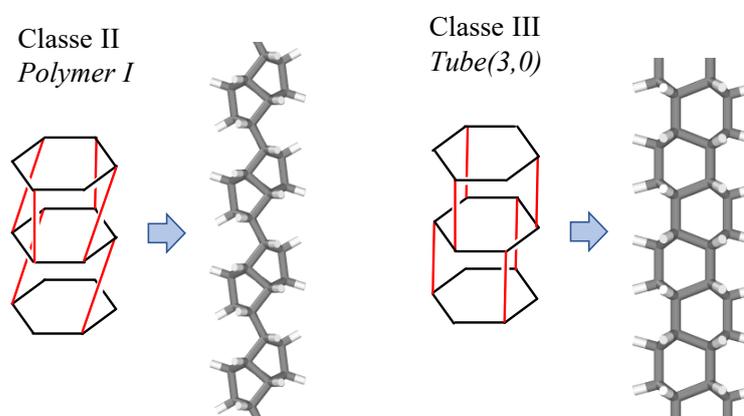


FIGURA 11 – Exemplos de NTHs de classe II e III.

A linguagem de programação Python foi escolhida por ter sido desenvolvida com a filosofia de diminuir o esforço e o tempo de trabalho do programador, priorizando a legibilidade do código e facilitando sua manutenção, que pode ser feita por cientistas e engenheiros, sem a necessidade de especialistas em programação [120]. Python é atualmente uma das linguagens mais utilizadas por programadores, com diversos pacotes voltados à ciência e à engenharia, como o módulo NumPy [121], que foi utilizado na implementação deste algoritmo genético. O algoritmo foi implementado como parte do pacote ChemStruct [122], que auxilia na preparação e análise estrutural de arquivos para simulação atômica.

A seguir (Seção 3.2.1) serão descritos os principais parâmetros de entrada ao algoritmo genético. As Seções subsequentes explicam em detalhes o seu funcionamento, cujas principais etapas são: criação de candidatos com genótipos coerentes (Seção 3.2.2); cálculo de propriedades desses candidatos através de um pacote de simulação escolhido pelo usuário (Seção 3.2.3); cálculo da probabilidade de seleção de cada candidato a partir do valor da propriedade fornecido pela simulação (Seção 3.2.4); e produção de uma nova geração através de operações como cruzamento e mutação (Seção 3.2.5). Por fim, as saídas do algoritmo genético são brevemente explicadas (Seção 3.2.6).

Por simplicidade, algumas partes da apresentação do algoritmo serão feitas utilizando-se o benzeno como molécula precursora padrão, uma vez que seus NTHs foram os primeiros descobertos e continuam sendo os mais pesquisados. No entanto, o algoritmo genético permite a utilização de qualquer molécula aromática. Algumas de suas ferramentas foram otimizadas para o tratamento de moléculas com anéis de cinco e seis membros (que representam a grande maioria dos NTHs sintetizados até então), em particular o identificador de ligações cruzadas; mas outras moléculas também podem ser utilizadas, ainda que com maior cuidado.

Após a exposição completa do algoritmo genético, é descrita a metodologia que foi empregada para testá-lo (Seção 3.2.7).

3.2.1 Entradas (*inputs*)

Os parâmetros dados como entrada ao algoritmo genético definem o problema de otimização a ser realizado, de maneira que é conveniente citar alguns deles antes de explicar o funcionamento do algoritmo em si. Esses parâmetros são divididos nos seguintes grupos: simulação para cálculo da aptidão, escolhas I/O (entrada e saída de dados), parâmetros de estrutura dos NTHs, parâmetros evolutivos principais, e parâmetros evolutivos secundários.

Os parâmetros relativos à simulação especificam qual programa será utilizado para o cálculo da aptidão (como um pacote de MD, DFT, etc), bem como o arquivo de entrada a ser dado a esse programa (cujo formato varia de acordo com o próprio programa). Também incluem o caminho para a pasta onde a evolução deve ser conduzida (onde os arquivos de saída do algoritmo serão escritos, e onde a simulação de cada NTH será realizada, em pastas separadas por candidato e por geração).

Os parâmetros I/O especificam se as pastas de simulação devem ser apagadas no

final da evolução, bem como se são desejados arquivos adicionais de saída, com informações suplementares sobre o processo evolutivo realizado.

Os parâmetros de estrutura dos NTHs especificam as restrições estruturais de sua evolução. O principal dado é um arquivo XYZ contendo a molécula aromática que deve ser usada como precursora na construção dos NTHs. Outras informações estruturais são o número de moléculas que devem ser usadas por NTH, e se os NTHs são periódicos ou não. Além disso, deve ser informada a classe dos NTHs, que precisa se manter fixa durante o processo evolutivo (NTHs de baixa energia podem ser divididos em duas classes, II e III, que possuem respectivamente 2:4 e 3:3 ligações com suas duas moléculas adjacentes [4]). Outros parâmetros de menor importância estão descritos nas próximas Seções, onde apropriado.

Os parâmetros evolutivos principais são o número de gerações e o número de candidatos por geração, bem como o número de candidatos replicados, cruzados (com e sem mutação), e mutados entre uma geração e outra.

Os parâmetros evolutivos secundários especificam quais metodologias devem ser utilizadas para o cálculo da probabilidade de um candidato ser selecionado para as operações evolutivas (replicação, cruzamento, etc). Isso inclui o equacionamento da aptidão (normalizada e tratada), da singularidade e da probabilidade – todas essas variáveis estão descritas na Seção 3.2.4, onde os diferentes métodos de cálculo são descritos.

3.2.2 Estratégia de criação dos genótipos

Como descrito na Seção anterior, os parâmetros de entrada do algoritmo incluem informações estruturais sobre os NTHs a serem estudados, como a molécula precursora e o número de moléculas por NTH. Logo, a única diferença entre os candidatos de um processo evolutivo é o padrão de ligações químicas formadas entre as moléculas precursoras, de maneira que esse padrão é uma escolha natural para o genótipo dos NTHs. Uma vez que este algoritmo genético foi implementado como parte do pacote ChemStruct, ele pode lidar facilmente com átomos e suas posições; no entanto, o genótipo em si não possui nenhuma informação sobre a posição dos átomos do NTH (ao contrário da maioria dos algoritmos genéticos utilizados em design de materiais, como discutido na Seção 2.3), apenas das ligações produzidas entre moléculas adjacentes. A representação que o algoritmo faz de um candidato está ilustrada na Figura 12, onde as ligações já foram formadas com base em um genótipo aleatório, mas os átomos continuam nas posições originais (relativas a um conjunto de moléculas empilhadas).

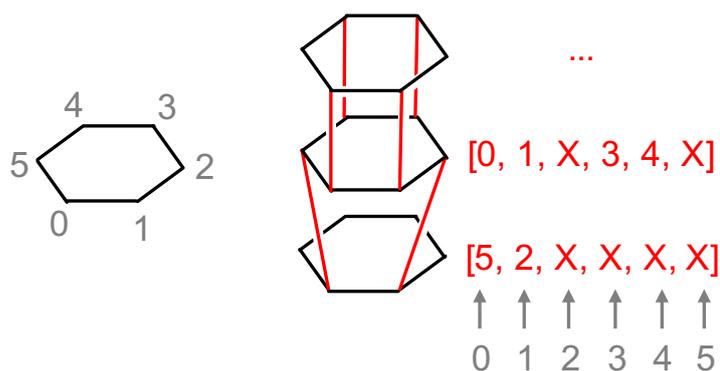


FIGURA 12 – Exemplo de código genético para um NTH de benzeno. Os números em cinza mostram a numeração das posições atômicas, feita em sentido anti-horário (positivo em torno do eixo z). As ligações estão representadas em vermelho, bem como o gene das moléculas correspondentes. Por exemplo, a primeira molécula (de baixo para cima) tem seu átomo de índice 0 ligado à posição 5; e seu átomo de índice 1 ligado à posição 2.

A Figura 12 também ilustra o genótipo utilizado pelo algoritmo genético, onde cada molécula possui uma lista de índices que informam, para cada um de seus átomos, com que átomo na molécula superior deve ser feita uma ligação. Essas listas são tratadas como genes, uma vez que são a menor parte do código genético que pode ser alterado de maneira independente (não se pode alterar um único índice sem ajustar os demais, uma vez que isso levaria a um átomo de carbono fazendo duas ligações com moléculas adjacentes, e portanto cinco ligações ao todo). O genótipo de um NTH é, portanto, uma lista de genes (uma lista de listas). Uma vez que um gene só contém informação a respeito de ligações com a molécula superior, a letra "X" é usada para marcar os átomos que se ligam com a molécula inferior.

A nomenclatura proposta por Xu et al. [4] para NTHs derivados de benzeno pode ser vista como uma versão condensada do método usado aqui, mas que só funciona para sistemas formados por duas moléculas. A notação usada neste algoritmo genético, por outro lado, pode ser estendida a NTHs de qualquer tamanho, como exigem os objetivos de sua implementação.

No início do processo evolutivo, os genótipos da primeira geração são gerados de maneira aleatória. Esse processo depende da classe à qual o NTH pertence. Por exemplo, o NTH mostrado na Figura 12 é de classe II, com 2 ligações com uma moléculas e $(6 - 2 =)$ 4 ligações com a outra. Em NTHs totalmente saturados, esse padrão se repete ao longo de toda a estrutura: pode haver alteração nas posições ligadas, mas não em seu número, uma vez que cada molécula de benzeno precisa fazer 6 ligações no total. Além disso, NTHs periódicos de classe II precisam ser constituídos de um número par de moléculas, para que suas extremidades sejam compatíveis.

Na criação de um genótipo aleatório, o algoritmo genético começa produzindo um gene para a primeira molécula (operação *start*). Ele estabelece duas ou três ligações com a molécula superior, conforme a classe informada pelo usuário (II ou III), e marca com "X" as demais posições.

Depois disso, o NTH passa por um processo de crescimento (operação *grow*), no qual as demais moléculas, em ordem ascendente, recebem genes que respeitam a molécula anterior; ou seja, os átomos que já estão comprometidos com ligações "para baixo" não podem se ligar "para cima", precisando ser marcados com "X". Por exemplo, no NTH da Figura 12 a primeira molécula está ligada às posições 5 e 2 da superior. Logo, a segunda precisa ter um "X" em suas posições 2 e 5, e índices numéricos em todas as outras. Portanto, a afirmação feita antes – de que um gene pode ser alterado de forma independente – não é rigorosa: ele pode ser editado, mas levando-se em conta as suas moléculas adjacentes, caso elas já tenham genes definidos. Uma vez que o crescimento dos NTHs é feito de baixo para cima, essa parte do algoritmo só precisa considerar o gene da molécula imediatamente inferior.

Após o crescimento, no caso de NTHs periódicos, as moléculas das pontas precisam ser unidas (operação *join*), isto é, ligadas uma à outra. Esse processo leva em conta a disponibilidade dos átomos de cada uma delas, ligando-os aleatoriamente. NTHs não-periódicos recebem átomos adicionais de hidrogênio nos carbonos subcoordenados de ambas as pontas.

A numeração dos átomos começa em 0 para manter o padrão utilizado em Python, e então segue ao longo do ciclo em sentido anti-horário (giro positivo em torno do eixo z), como mostrado na Figura 12. Além disso, a numeração deve ser feita de maneira consistente. No caso do benzeno, ela pode começar (índice 0) em qualquer carbono, uma vez que essa molécula é invariante com respeito a rotações discretas. No caso de outras moléculas aromáticas, entretanto, é importante escolher uma referência e utilizá-la de maneira consistente. Há certos casos para os quais o algoritmo faz a numeração dos átomos de maneira automática. Para derivados monossustituídos do benzeno, o carbono substituído recebe o índice 0. Para moléculas com um único heteroátomo, esse átomo marca o início da contagem: ele recebe o índice 0 caso possa fazer ligações (como o N da piridina), e, caso não possa (como o S do tiofeno), o índice 0 vai para o próximo átomo em sentido anti-horário. O usuário pode informar ao algoritmo quais espécies atômicas (do ciclo apenas) têm permissão para fazer ligações. Todas as moléculas monocíclicas já utilizadas na síntese de NTHs estão contempladas nessa numeração automática; entretanto, caso o usuário queira utilizar outro composto, ele pode fazer a numeração manualmente e fornecer o

arquivo XYZ de entrada da molécula. É preciso apenas garantir que os primeiros átomos listados sejam os ligantes, e que eles estejam ordenados conforme o ciclo (preferencialmente em sentido anti-horário).

Nem todos os genes gerados aleatoriamente são "aceitos" pelo algoritmo. O gene [1, 0, X, X, X, X], por exemplo, representa ligações cruzadas, que em nenhum caso poderiam ser estáveis. Genes com ligações cruzadas são descartados durante a criação de um genótipo, gerando-se novas listas aleatórias até que sejam obtidos índices aceitáveis. Esse método é simples e garante que os genes sejam produzidos sem um viés metodológico. Além disso, o custo computacional de gerar genes adicionais é desprezível, uma vez que é uma parcela relativamente pequena deles que resulta em ligações cruzadas (ca. 50 e 72 % em moléculas com quatro e seis átomos ligantes, respectivamente, enquanto apenas um valor superior a 99 % levaria a um custo computacional apreciável).

A identificação de ligações cruzadas é feita através da Tabela 2, onde os átomos i e j da molécula inferior estão ligados, respectivamente, aos átomos n e m da molécula superior. Uma vez que as moléculas em questão são cíclicas, é natural que a tabela apresente certas simetrias. A título de exemplo, a primeira combinação problemática apresentada na tabela deve ser lida da seguinte maneira: tomando-se como referência i e tendo-se $n = i$ e $j = i + 1$, o único valor que m não pode tomar é $m = j - 2$. Na próxima linha da tabela tem-se $n = i + 1$ e $j = i + 1$, caso em que os valores que m não pode tomar são $m = j - 1$ e $m = j - 2$. A existência desse método de eliminação prévia é um dos motivos pelos quais os átomos da molécula precursora devem ser numerados seguindo-se o ciclo aromático, uma vez que a identificação de ligações cruzadas foi construída com base nessa numeração. O filtro também retém casos em que um átomo está ligado à posição oposta a ele (0 e 3 do benzeno, por exemplo; ver Figura 12), uma vez que uma ligação desse tipo só é estável caso seja a única presente entre as duas moléculas, o que não ocorre em NTHs de classe II e III.

A rigor, a identificação de ligações cruzadas não é feita exatamente como apresentado na Tabela 2, que apresenta, em vez disso, o caso especial de moléculas com seis átomos ligantes. O algoritmo faz uso de uma versão generalizada dessa tabela. A notação $n = i$, por exemplo, significa que o átomo i está ligado "reto"/"de forma direta", ou seja, com o átomo de mesma posição na molécula superior. Similarmente, $n = i - 1$ e $n = i + 1$ significam que o átomo i está ligado "à esquerda" e "à direita", respectivamente. Qualquer outro valor ($n < i - 1$ ou $n > i + 1$) corresponde a átomos ligados "à esquerda distante" ou "à direita distante". Essas

i	$j = i + 1$	$j = i + 2$	$j = i + 3$
$n = i$	$m = j - 2$		
$n = i + 1$	$m = j - 1$	$m = j - 2$	
	$m = j - 2$		
$n = i + 2$	$m = j$	$m = j - 1$	$m = j - 2$
	$m = j - 1$	$m = j - 2$	
	$m = j - 2$		

TABELA 2 – Identificação de ligações cruzadas em moléculas com seis átomos ligantes. Sejam i e j índices de átomos na molécula inferior que estejam ligados, respectivamente, aos átomos de índice n e m na molécula superior. As combinações apresentadas na tabela representam ligações cruzadas.

contagens são feitas até o átomo ligado "à posição oposta", sendo esta a última classificação de ligação, usada apenas para um número par de átomos ligantes. O identificador também leva em conta a natureza cíclica das moléculas: por exemplo, no caso do benzeno com uma ligação entre $i = 5$ e $n = i + 1 = 6$ (isto é, com ligação "à direita"), ocorre a normalização para $n = 0$, o que permite a identificação correta do átomo (que estariam indexados de 0 a 5). A Tabela 3 é a versão generalizada da Tabela 2. A título de ilustração, a Figura 13 mostra a classificação das posições em moléculas representativas.

i	$j = i + 1$	$j = i + 2$	$j = i + 3$
reto	à esquerda distante		
à direita	à esquerda	à esquerda distante	
	à esquerda distante		
à direita distante	reto	à esquerda	à esquerda distante
	à esquerda	à esquerda distante	
	à esquerda distante		

TABELA 3 – Identificação de ligações cruzadas em qualquer molécula. Sejam i e j índices de átomos na molécula inferior. As combinações apresentadas na tabela representam ligações cruzadas.

Além de ligações aleatórias, o algoritmo também pode produzir rotações discretas nas moléculas, caso o usuário especifique. Essas rotações são representadas como o primeiro índice de cada gene, antes da representação das ligações (Figura 14). Isso não é necessário no caso do benzeno, mas adiciona um grau de liberdade importante aos NTHs derivados de moléculas com menor número de graus de simetria, como piridina e tiofeno. Sendo b o número de átomos ligantes por molécula, as rotações são feitas em múltiplos de $360^\circ/b$, o que facilita o alinhamento dos átomos na maioria dos casos. A única exceção são moléculas com átomos não-ligantes no ciclo. O tiofeno, por exemplo, possui um anel de cinco átomos, dos quais apenas quatro (os carbonos) podem se ligar com as moléculas adjacentes. Como resultado, ele tem quatro rotações

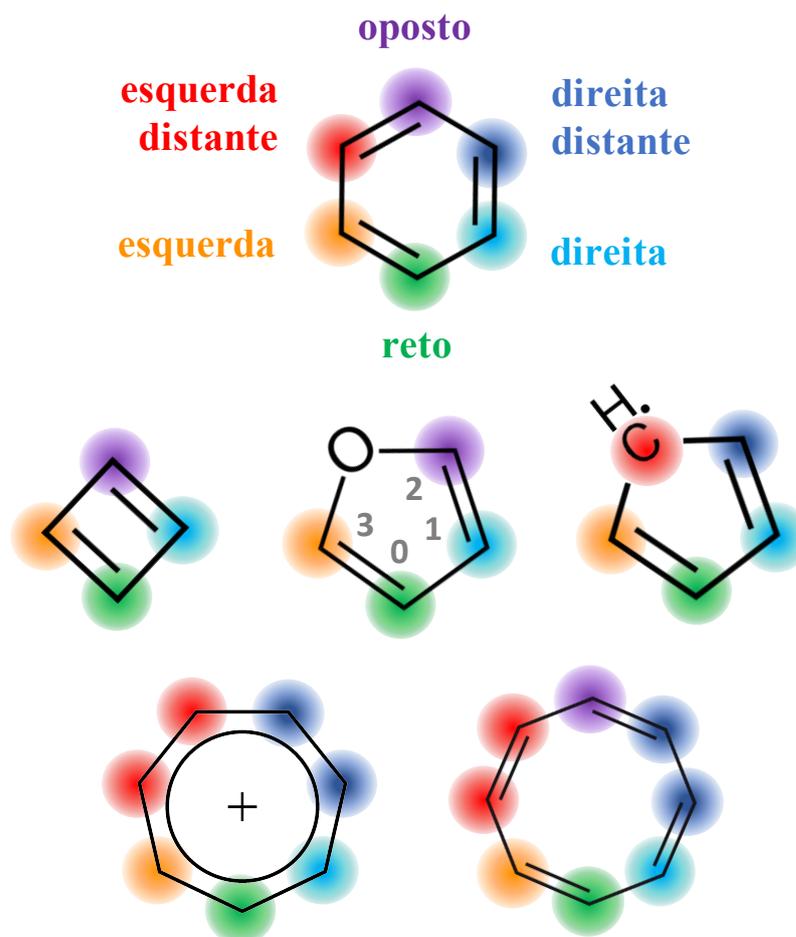


FIGURA 13 – Classificação de ligações em moléculas representativas. A legenda das cores está indicada no benzeno, acima. Em todos os casos o átomo inferior é tomado como referência. A numeração dos átomos ligantes (carbonos) do furano está mostrada em cinza.

possíveis (0° , 90° , 180° e 270°) em vez de cinco, o que não garante um alinhamento perfeito das moléculas. No entanto, isso não é um problema, uma vez que são as ligações químicas, e não as posições atômicas, que o genótipo codifica. (A próxima seção discute uma possível pré-otimização das posições atômicas, realizada antes da simulação e do cálculo de aptidão.) Os índices dos átomos são rotacionados junto à molécula, de maneira a manter coerente o método de eliminação de ligações cruzadas; isto é, os átomos de mesmo índice estão sempre (aproximadamente) alinhados, independentemente da rotação das moléculas. Logo, dentro do algoritmo genético, os índices não representam átomos, mas posições.

Uma propriedade fundamental ao funcionamento de algoritmos genéticos é que as alterações no valor da aptidão de um candidato sejam aproximadamente proporcionais às alterações realizadas em seu genótipo, de maneira que uma pequena mutação em um candidato apto sempre gere outro candidato apto [111]. O genoma escolhido para os NTHs garante essa

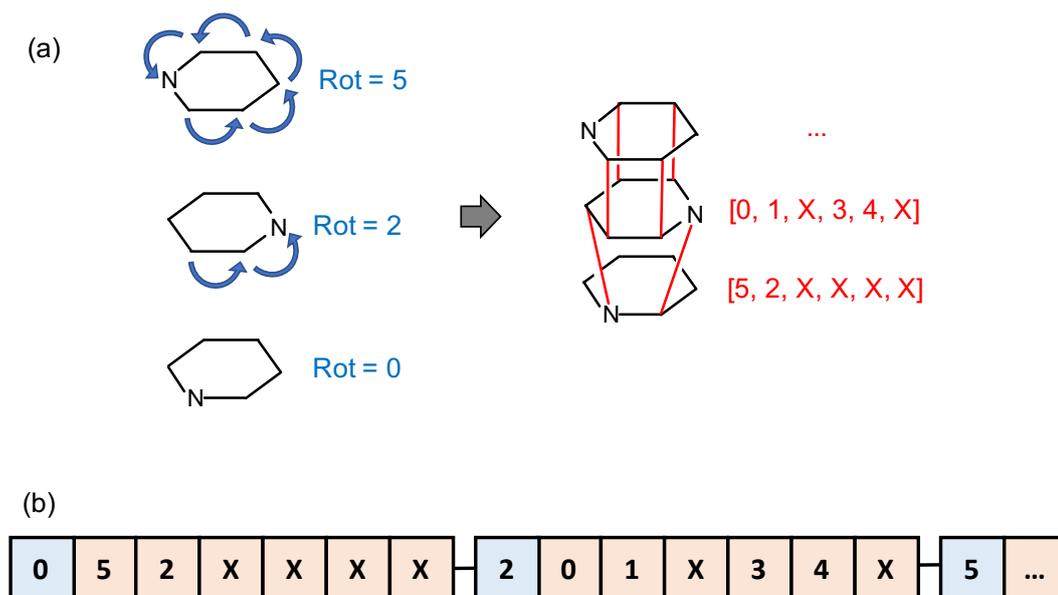


FIGURA 14 – Representação completa do genótipo, incluindo índices de rotação. (a) Rotações e ligações entre moléculas de piridina. (b) Genótipo correspondente.

propriedade, uma vez que pequenas alterações genéticas causarão pequenas mudanças estruturais, que por sua vez produzirão (em geral) pequenas variações nas propriedades do material. Essa proporcionalidade deverá ser ainda maior para NTHs não-periódicos, nos quais as partes da estrutura (cada uma codificada em um gene) são mais independentes umas das outras, podendo ter sua estrutura mais facilmente relaxada.

3.2.3 Simulação e Cálculo de Aptidão Bruta

A aptidão bruta (*raw fitness*) recebe esse nome para diferenciá-la da aptidão (normalizada e tratada), que será definida na próxima Seção. O cálculo da aptidão bruta é inteiramente customizável pelo usuário. Nesta etapa, a estrutura do NTH é dada como entrada a um programa de simulação computacional, que gera como saída o valor da aptidão bruta – isto é, qualquer variável definida pelo usuário, em geral alguma propriedade calculada do sistema (também é possível defini-la como uma variável composta, que leve em conta duas ou mais propriedades ponderadas).

A princípio, qualquer método pode ser utilizado para o cálculo das propriedades do sistema (simulações de MD clássica, cálculos DFT, etc). O pacote ChemStruct, dentro do qual o algoritmo genético foi implementado, já possui classes que ajudam a escrever inputs para os pacotes LAMMPS e Quantum Espresso (softwares de MD clássica e DFT, respectivamente). Além disso, qualquer outro programa pode ser utilizado, contanto que o algoritmo genético seja

alterado para escrever os arquivos nos formatos necessários.

Apesar da flexibilidade do algoritmo genético, há alguns cuidados que devem ser tomados durante sua utilização. O mais importante diz respeito ao tipo de entrada que o pacote de simulação escolhido aceita para estruturas atômicas. Em muitos casos (como na maioria dos programas de DFT e MD clássica reativa), as únicas informações estruturais dadas como entrada são a espécie e a posição dos átomos. Outras estruturas topológicas, como ligações químicas, são inferidas pelo próprio método de simulação, em geral a partir da distância entre os átomos. Logo, as posições atômicas fornecidas ao programa de simulação devem estar pelo menos próximas às da estrutura otimizada, de maneira que a topologia correta possa ser inferida. Como descrito na Seção anterior, o genótipo dos NTHs só codifica ligações químicas, e a representação que o ChemStruct faz de cada candidato possui posições atômicas que não correspondem à estrutura desejada. Logo, é necessário realizar uma relaxação estrutural antes do cálculo da aptidão bruta, como ilustrado na Figura 15. Essa pré-otimização deve ser feita com algum método que respeite o genótipo, isto é, que aceite uma lista de ligações como entrada da simulação. Uma opção são simulações de MD empregando um potencial não-reativo no programa LAMMPS, que podem ser realizadas com diversos potenciais clássicos e suas parametrizações. (Um exemplo de pré-otimização é descrito na Seção 3.2.7.)

A pré-otimização tem como função principal obter posições atômicas que de fato representam o NTH, e que podem então ser fornecidas à simulação principal para o cálculo da aptidão bruta. Isso é importante porque, se a estrutura mostrada à esquerda da Figura 15 fosse fornecida a um cálculo de DFT, por exemplo, o sistema seria tratado como um conjunto de moléculas de benzeno empilhadas, sem ligações entre si. Além disso, mesmo que a simulação principal aceite uma lista de ligações como parte do input, a pré-otimização ajuda a reduzir o custo computacional total, uma vez que obtém posições atômicas aproximadas utilizando um campo de força mais simples. Em particular, quando a simulação principal não envolve uma segunda (e mais precisa) etapa de relaxação, a pré-otimização é indispensável – mas nesses casos é preciso certificar-se de que os parâmetros utilizados na pré-relaxação são bons o bastante para o cálculo de aptidão bruta desejado.

A pré-otimização tem uma segunda utilidade: ela permite a comparação das ligações químicas do NTH relaxado com aquelas postuladas pelo seu genótipo, encontrando estruturas instáveis e atuando como um segundo filtro de candidatos. O algoritmo genético foi implementado como parte do pacote ChemStruct, cujas funcionalidades permitem a listagem de ligações

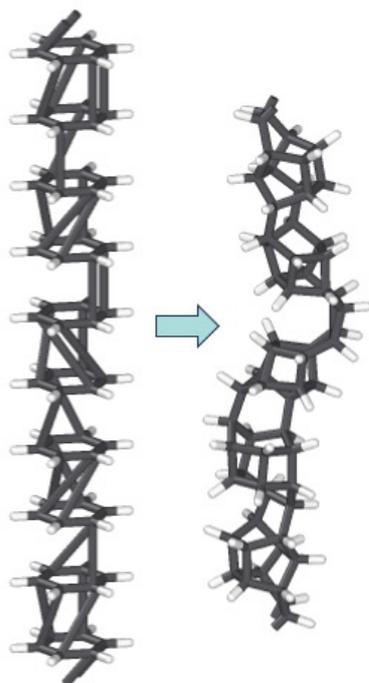


FIGURA 15 – Pré-otimização estrutural de um NTH periódico. A estrutura à esquerda mostra as posições atômicas das moléculas precursoras, bem como as ligações geradas entre elas pelo algoritmo genético. Na estrutura à direita, as coordenadas dos átomos correspondem ao mínimo local de energia dessa estrutura.

químicas com base em distâncias interatômicas. Logo, é possível fazer um pós-tratamento no sistema relaxado (dado como saída do programa de simulação escolhido) e identificar quais das ligações químicas indicadas no genótipo se mantiveram após a pré-otimização. Se pelo menos uma das ligações não for computada pelo ChemStruct (ou se for encontrada uma adicional), sabe-se que seus átomos estão separados demais (ou próximos demais), o que significa que o NTH em questão é instável. Mesmo que a simulação principal conseguisse encontrar um mínimo local de energia para a estrutura, quebrando certas ligações e produzindo novas, o novo material não seria aquele codificado pelo genótipo, de maneira que a aptidão calculada não teria nenhum significado. Em casos extremos, até mesmo as ligações internas das moléculas precursoras poderiam ser quebradas, o que desclassificaria a estrutura final como um possível NTH. Esse segundo filtro atua dando aos candidatos instáveis uma aptidão bruta nula, que produzirá uma probabilidade zero de serem selecionados nas operações evolutivas.

É importante ressaltar que esta otimização local ocorre em uma paisagem de aptidão diferente da procura global do algoritmo genético. O algoritmo produz variações de genótipo (rotações e ligações entre moléculas) para maximizar a aptidão (definida pelo usuário). Já a otimização local varia posições atômicas (e o tamanho L do sistema, no caso de NTHs periódicos)

para minimizar a energia potencial. Essa relaxação estrutural é feita através de uma combinação do método de gradiente conjugado (como implementado na versão de Polak-Ribiere no software LAMMPS, chamada no comando *minimize*) e relaxação em *ensemble* NPT a uma temperatura de 0.1K.

A implementação padrão do algoritmo genético procura pelo valor da aptidão bruta em um arquivo de texto chamado *fitness*, que deve ser gerado pela simulação principal; mas o local onde esse valor é procurado pelo algoritmo pode ser alterado pelo usuário, podendo fazer parte dos arquivos de saída do programa de simulação. É importante notar que o algoritmo genético tentará gerar NTHs que maximizam a aptidão bruta; logo, caso se deseje minimizar alguma propriedade (como a energia do sistema), é preciso multiplicar a propriedade de interesse por -1 na definição da aptidão bruta.

3.2.4 Cálculo de Aptidão, Similaridade, Singularidade e Probabilidade de seleção dos candidatos

A aptidão (*fitness*) F_i do candidato i é calculada a partir de sua aptidão bruta (*raw fitness*) f_i , segundo a Equação 3.1, onde f_{min} e f_{max} são os valores mínimo e máximo de aptidão bruta encontrados entre os candidatos daquela geração, excetuando-se os valores nulos (relativos a NTHs instáveis).

$$F_i = \frac{1}{2} \left[1 + \tanh \left(2 \frac{f_i - f_{min}}{f_{max} - f_{min}} - 1 \right) \right] \quad (3.1)$$

Essa definição tem como principal objetivo normalizar a aptidão de cada candidato em relação à faixa da valores de sua geração, uma vez que é com essa população que ele está competindo [99]. Além disso, a Equação 3.1 garante que F_i esteja entre 0 e 1, antevendo seu uso como probabilidade. Mais precisamente, os valores de F_i caem no intervalo entre ca. 0.12 e 0.88, o que faz com que até o candidato menos apto tenha alguma chance de seleção, e que nem o mais apto tenha chance de 100%.

A probabilidade P_i de o candidato i ser escolhido para ter seus genes carregados para a próxima geração (através das operações descritas na Seção 3.2.5) pode ser definida de diversas maneiras, segundo a escolha do usuário. A opção mais simples é $P_i = F_i$, na qual a aptidão é o único fator de escolha.

No entanto, algoritmos genéticos se beneficiam da utilização de um segundo fator: a

singularidade do candidato, que quantifica o quão diferente ele é dos demais. Incluir o valor de singularidade no cálculo da probabilidade P_i garante que o algoritmo produzirá um conjunto mais diverso de materiais, reduzindo a chance de ele ficar preso em máximos locais de aptidão [99]. Em alguns casos relatados na literatura, a definição da singularidade é simples: conta-se quantas vezes o material apareceu nas gerações anteriores, então inverte-se esse valor com alguma equação apropriada. Entretanto, isso ainda permite que o algoritmo se mova por tempo demais nas proximidades de um máximo local, produzindo mutações e gerando candidatos muito parecidos entre si. Portanto, neste estudo sobre NTHs, foi utilizada uma definição de singularidade que não apenas considera candidatos iguais, mas que também pondera aqueles cujo genótipo é parecido.

Primeiramente, é preciso quantificar o quão parecidos são dois NTHs – digamos i e j . Isso é feito através de uma variável S_{ij} , definida pela Equação 3.2, onde s_{ij} é uma contagem de similaridade entre seus genótipos, e s_{ij}^{max} é a maior contagem possível (obtida caso os dois NTHs sejam iguais). Logo, S_{ij} é a fração do genótipo i que é igual ao genótipo j .

$$S_{ij} = \frac{s_{ij}}{s_{ij}^{max}} \quad (3.2)$$

A contagem s pode ser feita de duas maneiras, segundo a especificação do usuário. A primeira é contando-se cada índice de cada gene, ou seja, cada número de ligação e de rotação. A segunda é contando-se apenas genes iguais, ou seja, quando todos os índices (rotação e ligações) são os mesmos em ambos os genótipos. Em ambos os métodos, a última molécula é ignorada no caso de NTHs não-periódicos, uma vez que seu gene é fixo (não codifica nenhuma ligação). Além disso, no caso em que os índices são contados individualmente, se os NTHs não incluírem rotação de moléculas, então os índices de rotação são ignorados, tanto em NTHs periódicos quanto não-periódicos. Em todos os casos, a contagem s_{ij} é feita para todas as possíveis rotações do NTH j como um todo (bem como translações, no caso periódico), considerando-se assim as principais operações de simetria; e é utilizada a contagem mais alta encontrada.

O fator S_{ij} é calculado para todos os pares de candidatos. Uma vez que isso é feito, pode-se computar a similaridade (*similarity*) S_i de cada candidato através da Equação 3.3, onde c e k são parâmetros fornecidos pelo usuário (com valores padrão de 1).

$$S_i = \sum_j \frac{S_{ij}^c}{a} \quad (3.3)$$

Uma vez que cada S_{ij} tem um valor entre 0 e 1, a utilização de um c alto tem como função reduzir a contribuição de pares parcialmente similares de NTHs. Por exemplo, com $c = 1$ o peso de um NTH igual a i é igual ao peso de dois NTHs com metade do genótipo igual ao de i ; mas com $c > 1$ o segundo cenário leva a uma similaridade S_i menor, o que pode ser desejável. Já o valor de a pode ser utilizado para normalizar S_i por algum número, como o número de candidatos por geração, o que permite a comparação de evoluções com parâmetros diferentes. Por fim, o somatório da Equação 3.3 pode ser feito entre todos os candidatos da geração atual, ou entre todas as gerações simuladas até então, segundo a escolha do usuário.

A singularidade (*uniqueness*) U_i de cada candidato é calculada a partir de sua similaridade S_i , segundo a Equação 3.4 [99]. Essa expressão garante que $0 < U_i < 1$, já prevendo o uso da singularidade no cálculo de probabilidade.

$$U_i = \frac{1}{\sqrt{1 + S_i}} \quad (3.4)$$

Como mencionado anteriormente, a probabilidade de seleção P_i de cada candidato pode ser definida de diferentes maneiras, segundo a escolha do usuário. A definição mais completa é $P_i = F_i U_i$. Também podem ser utilizadas as opções $P_i = F_i$ e $P_i = U_i$, além de $P_i = kF_i + (1 - k)U_i$, onde k é um valor entre 0 e 1 dado pelo usuário.

Por fim, uma nota sobre a implementação da classe que representa os NTHs, a *Nanoread()*. A igualdade entre dois objetos dessa classe, implementada através do método mágico `__eq__()`, foi definida como dois NTHs cujos genótipos tenham $S_{ij} = 1$. Essa verificação de igualdade é utilizada diversas vezes pelo algoritmo; por exemplo, quando é preciso garantir que não haja NTHs repetidos em uma mesma geração.

3.2.5 Produção de uma nova geração: Replicação, Mutação, Cruzamento e Aleatorização

Uma vez computada a probabilidade P_i de cada candidato da geração atual, esses valores são usados para pesar sua chance de seleção durante a criação da próxima geração através das operações evolutivas: replicação, mutação, cruzamento (simples e inteligente) e re-aleatorização – *replication*, *mutation*, *crossover* e *rerandom*, respectivamente. O usuário escolhe quantos novos candidatos serão gerados a partir de cada um desses métodos.

A replicação consiste em criar um novo candidato com o mesmo genótipo daquele selecionado. NTHs já simulados não são testados de novo; em vez disso, seu valor de aptidão

bruta é lido de um dicionário cuja chave é seu código genético (sendo contempladas as operações de simetria, isto é, translação e rotação do genótipo). A replicação tem como objetivo perpetuar os genes mais aptos, um processo também chamado de elitismo.

A mutação consiste em realizar a menor alteração possível no genótipo de um candidato, o que no caso de NTHs significa alterar um de seus genes – isto é, as ligações entre duas moléculas subsequentes. Essa edição não é totalmente independente do restante do código genético, uma vez que, se for feita às cegas, pode produzir átomos com ligações demais, como discutido na Seção 3.2.2. Logo, na prática a mutação "embaralha" os índices dos átomos que já estavam ligados, criando novas ligações (com o método *join*). Como em todas as operações além da replicação, o novo candidato só é aceito caso seja diferente do NTH que lhe deu origem. Em todas as operações, incluindo a replicação, ele deve ser diferente de todos os outros candidatos já admitidos para a próxima geração, bem como respeitar a análise de ligações cruzadas discutida na Seção 3.2.2. Essas limitações fazem com que muitas mutações sejam descartadas, e o algoritmo continua gerando-as até que o número de mutações especificado pelo usuário seja atingido. O usuário também pode alterar o número máximo de genes alterados na mutação, cujo valor padrão é 1.

O cruzamento simples consiste em selecionar dois candidatos parentais e "cruzar" seus genótipos para produzir dois NTHs filhos. Ambos os NTHs filhos são adicionados à nova geração – a menos que isso ultrapasse o número de cruzamentos escolhido pelo usuário, caso no qual apenas um deles é adicionado. O ponto de corte do código genético é escolhido aleatoriamente, e o gene que conecta as duas partes é editado (com o método *join*) de maneira a compatibilizá-las. Caso o NTH seja periódico, o último gene (que codifica as ligações entre a última e a primeira molécula) também é adaptado. Em função dessas compatibilizações, o cruzamento simples entre NTHs envolve, intrinsecamente, uma ou duas mutações (no caso não-periódico e periódico, respectivamente). Logo, para um NTH periódico com oito moléculas, um cruzamento simples produz novos candidatos com seis dos oito genes parentais (em qualquer proporção) e dois genes aleatórios novos.

Também é possível fazer um cruzamento inteligente, no qual é procurado um ponto de compatibilidade entre os dois genótipos – isto é, um ponto no qual os dois genes são iguais. Essa procura é feita rotacionando (e translacionando, caso o NTH seja periódico) um dos genótipos. É procurado um gene igual caso o NTH seja não-periódico, e dois caso ele seja periódico. No cruzamento inteligente, os dois pontos de corte são variáveis, enquanto que no cruzamento

simples um deles sempre ocorre na fronteira periódica (entre a última e a primeira molécula). O usuário pode especificar quantos cruzamentos de cada tipo devem ser feitos. A utilização do cruzamento inteligente pode beneficiar NTHs parecidos, gerando um viés evolutivo, mas também garante que nem todos os cruzamentos ocorram com mutação, o que melhora o desempenho de algoritmos genéticos ao diminuir o "ruído aleatório" dos novos candidatos [99].

Por fim, a re-aleatorização consiste em escolher um ponto aleatório do genótipo e fazer o NTH crescer de novo a partir dali (pelo método *grow*). Como resultado, uma fração (de tamanho aleatório) do material é alterada aleatoriamente. O último gene é adaptado (com o método *join*), o que garante que as duas pontas do NTH sejam compatíveis. Esta é a última operação realizada durante a criação de uma nova geração. Candidatos são selecionados para essa operação até que a nova geração tenha atingido o número total de candidatos. O objetivo desta operação é adicionar uma quantidade apreciável de genes novos ao processo evolutivo; mas, uma vez que as outras operações fazem isso em menor escala, a re-aleatorização pode ser ignorada pelo usuário.

3.2.6 Saídas (*outputs*)

Diversas informações sobre o processo evolutivo são fornecidas na forma de arquivos de saída (*outputs*), em maior ou menor quantidade, conforme a escolha do usuário. Por exemplo, o arquivo *precursor.xyz* contém as coordenadas atômicas dos átomos na ordem utilizada pelo algoritmo genético – ou seja, seguindo o padrão discutido na Seção 3.2.2, e não necessariamente na ordem do arquivo de entrada dado pelo usuário.

Também é fornecida uma tabela com os valores de aptidão bruta computados em cada geração, uma vez que, em geral, esses valores representam a propriedade de interesse. Além disso, opcionalmente, são fornecidas tabelas com os valores de aptidão F_i , similaridade S_i , singularidade U_i e probabilidade P_i , cuja principal função é ajudar o usuário a avaliar a performance do algoritmo genético dentro dos parâmetros fornecidos. A memória de aptidão bruta, por sua vez, é uma lista que relaciona o código genético de todos os candidatos avaliados com sua respectiva aptidão bruta, em ordem de simulação. Uma lista de aptidão bruta média por geração ajuda a estimar a performance do algoritmo genético. É fornecida a média total, a média sem as aptidões nulas, e a média dos primeiros N candidatos mais aptos, onde N é o número de candidatos produzidos por replicação, cruzamento e mutação – descartando-se, assim, os candidatos re-aleatorizados.

Além disso, também ficam disponíveis as pastas onde as simulações de pré-otimização

e simulação principal são realizadas, com arquivos XYZ que permitem uma rápida visualização das estruturas sendo estudadas. Caso o usuário especifique, essas pastas são apagadas no final do processo evolutivo (o que é útil quando são realizados muitas evoluções independentes).

3.2.7 Teste do algoritmo genético

O algoritmo genético foi testado para NTHs de classe II derivados de benzeno, utilizando-se uma medida de energia potencial como aptidão, de maneira a encontrar os candidatos com maior estabilidade relativa. Foram simulados NTHs não-periódicos, uma vez que seu maior número de graus de liberdade na relaxação estrutural garante maior independência aos efeitos de genes individuais na aptidão, o que melhora a performance de algoritmos genéticos [95]. Foram testados NTHs de diferentes tamanhos, formados por 4, 8 e 16 moléculas. Para cada caso foram rodadas 100 evoluções independentes, cada uma com 20 gerações de 20 candidatos, totalizando 400 NTHs por evolução e 40000 para cada tamanho (valores típicos em trabalhos sobre algoritmos genéticos aplicados à otimização de materiais [99]). Para comparação, também foram gerados e simulados 40000 NTHs aleatórios de cada tamanho (aos quais foi aplicado o filtro que retém ligações cruzadas). Nos processos evolutivos, dos 20 candidatos das novas gerações, 4 foram replicados estocasticamente da geração anterior e os outros 16 foram cruzados: 8 com cruzamento simples, que envolve uma mutação intrínseca, e 8 com o cruzamento inteligente, no qual não há mutação (para garantir que não haja aleatoriedade demais no processo evolutivo [99]). Foi utilizada a opção de similaridade que compara genes inteiros entre todos os candidatos do processo evolutivo até então, e a probabilidade de seleção foi calculada como o produto entre a aptidão tratada e a singularidade.

A aptidão bruta (energia potencial do sistema) foi calculada através de simulações de MD clássica, mais especificamente uma relaxação estrutural (em *ensemble* NPT) com o software *open-source* LAMMPS [123]. No LAMMPS, simulações não-reativas permitem (e exigem) uma lista de ligações químicas como entrada, de maneira que o padrão de ligações no material seja mantido durante a simulação mesmo que as posições iniciais dos átomos estejam muito longe do mínimo de energia. Simulações reativas, por outro lado, não permitem que as ligações sejam dadas como entrada, de maneira que as coordenadas atômicas iniciais já devem estar próximas da estrutura desejada. Uma vez que o algoritmo genético não altera a posição dos átomos, fazendo apenas uma lista das ligações geradas entre as moléculas, a relaxação dos NTHs foi feita em duas etapas: uma primeira etapa com MD não-reativa, para "corrigir" as coordenadas atômicas, e

uma segunda etapa reativa, para calcular com maior precisão a energia potencial.

A primeira etapa da relaxação estrutural foi feita com parâmetros aproximados, tomados de uma parametrização do CHARMM para moléculas orgânicas em geral (*CHARMM General Force Field*, CGenFF [124]). Uma vez que esta etapa é apenas uma pré-relaxação, a precisão dos parâmetros não é de grande importância, e qualquer outro potencial não-reativo poderia ter sido usado. A segunda etapa da relaxação utilizou o potencial ReaxFF [78] com uma parametrização otimizada para interações C/Ni e grafeno fluoretado [88, 125], chamada daqui em diante de ReaxFC. A parametrização ReaxFC permitiu uma previsão razoável das propriedades mecânicas de NTHs de benzeno a um baixo custo computacional (comparado a cálculos de DFT) [2]. Uma verificação adicional foi feita aqui para o cálculo da energia dos NTHs estudados por Xu et al. [4], demonstrando que o ReaxFC faz boas previsões para estabilidade relativa de isômeros de NTH derivados de benzeno, apresentando as mesmas tendências gerais dos cálculos de DFT (Figura 16). A única grande exceção (tanto no caso da energia quanto das propriedades mecânicas) é o primeiro NTH mostrado na Figura 16, correspondente à estrutura *Tube(3,0)*, cuja estabilidade e resistência mecânica são superestimadas pelo ReaxFC (possivelmente devido à similaridade que sua estrutura tem com o diamante).

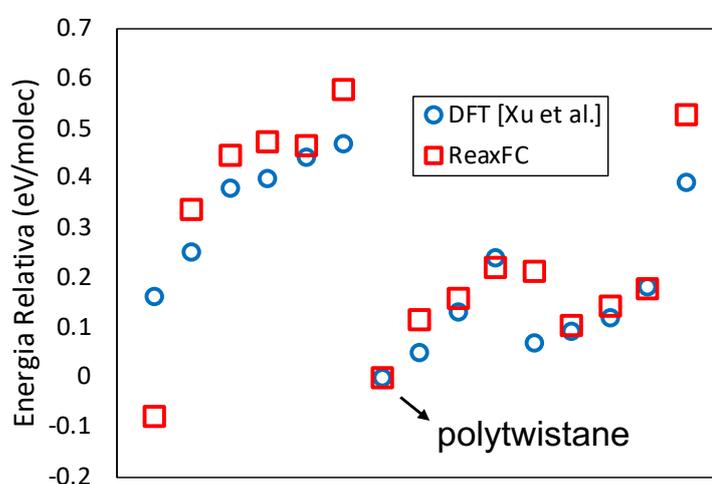


FIGURA 16 – Comparação da energia prevista por DFT e pelo ReaxFC para diferentes NTHs derivados de benzeno. A energia do *polytwistane* foi tomada como referência, uma vez que esse é o NTH mais estável segundo DFT [4].

Utilizando-se a energia potencial como aptidão bruta, uma evolução com 20 gerações de 20 candidatos leva cerca de uma hora (em um único processador, sem paralelização), ou seja, uma média de nove segundos por candidato. Esse tempo é majoritariamente devido à simulação de MD reativa.

Também foram feitas evoluções em menor escala utilizando-se o Módulo de Young como aptidão bruta, com o objetivo de demonstrar que o algoritmo pode ser aplicado ao design de NTHs com diferentes propriedades. Foi utilizada a mesma metodologia descrita anteriormente para o caso da energia potencial, com exceção de que foram usados NTHs periódicos. O teste simulado de tração uniaxial foi realizado em acordo com trabalhos semelhantes sobre NTHs [2, 42], com uma taxa de deformação aplicada de $1 \times 10^{-3} \text{ ps}^{-1}$ até que se atinja uma deformação de 3%. O Módulo de Young foi calculado como a inclinação da curva de tensão-deformação até esse ponto. Nos resultados está mostrado o Módulo de Young 1D, em unidades de força (nN), uma vez que a versão 3D dessa propriedade exige a definição da área da seção transversal do material, que não está bem definida no caso de materiais 1D ultrafinos, como NTHs. Essa representação vem sendo empregada em outros trabalhos teóricos voltados à análise das propriedades mecânicas desses materiais [2, 9, 19, 42].

O cálculo da tensão é feito tanto para fins de relaxação estrutural quanto para cálculo das curvas de tensão-deformação. Ele é feito pelo comando *compute stress/atom* do LAMMPS. A componente n, m do tensor de tensão 1D τ_{nm}^i (em unidades de força) do átomo i é calculado através da expressão virial mostrada na Equação 3.5 [76, 77], onde L é o comprimento do NTH, M_i é a massa do átomo, v_k^i é a componente k da velocidade do átomo, e r_m^{ij} e F_n^{ij} são a distância e a força entre o átomo e um vizinho j . (A tensão 3D tem um cálculo análogo, mas utilizando-se o volume V do sistema no lugar do comprimento L no NTH.)

$$\tau_{nm}^i = \frac{1}{L} \left[-M_i v_m^i v_n^i + \frac{1}{2} \sum_{j \neq i} r_m^{ij} F_n^{ij} \right] \quad (3.5)$$

4 RESULTADOS E DISCUSSÃO

4.1 CÁLCULOS DE DFT DAS PROPRIEDADES DE NANOFILAMENTOS DE CARBONO DERIVADOS DE MOLÉCULAS HETEROCÍCLICAS DE CINCO MEMBROS

4.1.1 Estrutura Atômica

A síntese de NTHs a partir de furano [7] e tiofeno [8] foi reportada pela primeira vez em 2019. As amostras foram caracterizadas via difração de raios X, cujos resultados foram comparados com padrões simulados a partir de cristais de diferentes estruturas. Essa comparação sugere fortemente que NTHs derivados de moléculas de cinco membros são bem representados pelas configurações *syn*, *anti* e *syn-anti*. Essas estruturas estão mostradas na Figura 17 para o caso do tiofeno; os materiais derivados do furano são equivalentes, exceto pela presença de átomos de oxigênio no lugar do enxofre. Outras moléculas de cinco membros também podem (hipoteticamente) produzir esses NTHs, nos quais os heteroátomos são substituídos pelo grupo funcional da molécula precursora (por exemplo, NH no caso do pirrol).

Essas três configurações são resultado de uma série de cicloadições [4 + 2], nas quais todos os quatro carbonos das moléculas formam ligações e se tornam saturados. Estruturas parecidas foram encontradas em materiais derivados de benzeno e piridina, nas quais os heteroátomos são substituídos por ligações duplas C=C ou C=N [5, 10, 11]. No caso de moléculas de seis membros, entretanto, o maior número de átomos aptos a reagir gera materiais com estrutura mais complexa: os átomos envolvidos nas ligações duplas podem reagir para formar novas configurações, e outras rotas iniciais (que não a das cicloadições) podem produzir estruturas diferentes.

As três configurações apresentadas na Figura 17 se diferenciam pela distribuição dos heteroátomos ("apêndices" ou "grupos anexados") ao longo do filamento. A configuração *syn* tem todos os apêndices do mesmo lado, enquanto a *anti* os tem de maneira alternada. A estrutura *syn-anti* é um meio-termo: o apêndice aparece duas vezes do mesmo lado antes de aparecer do outro. Uma maneira alternativa de classificar materiais dessa natureza é pelo número de vezes seguidas que o heteroátomo aparece do mesmo lado da cadeia principal. Chamando esse número de n , podemos classificar a estrutura *anti* como tendo $n = 1$, e a *syn-anti* como tendo $n = 2$. Outros valores finitos de n são teoricamente possíveis, mas desnecessários para modelar

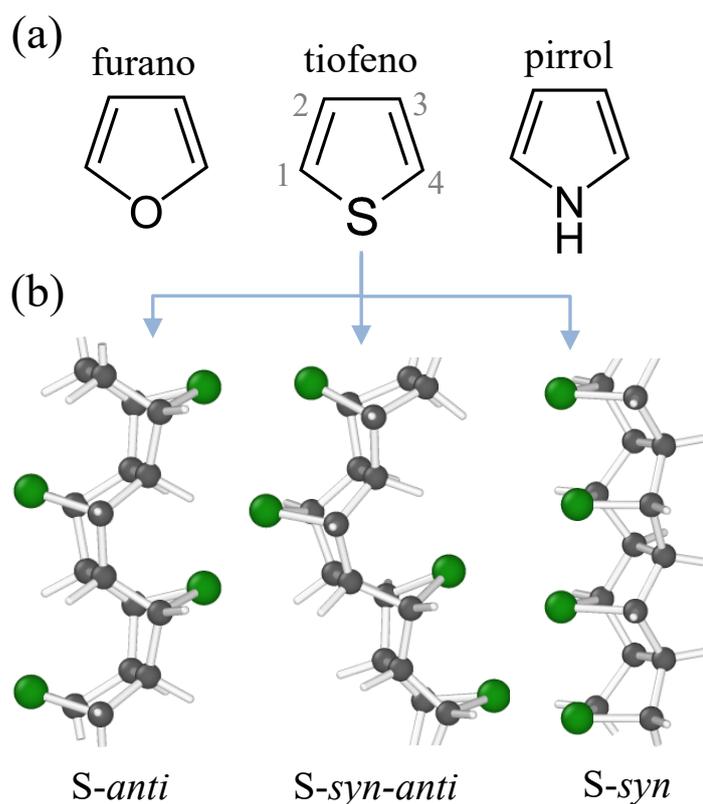


FIGURA 17 – Estrutura dos NTHs estudados. (a) Estrutura química das moléculas precursoras. (b) Estrutura atômica dos NTHs derivados do tiofeno. Reproduzido da referência [12] com permissão da Royal Society of Chemistry / PCCP Owner Societies.

o material real. A configuração *syn* é o caso extremo, no qual $n = \infty$.

Foram então estudadas as estruturas *syn*, *anti* e *syn-anti* derivadas de diferentes moléculas aromáticas heterocíclicas de cinco membros: tiofeno, furano, pirrol e fosfol. As duas primeiras moléculas já foram utilizadas em experimentos conforme mencionado acima [7, 8], e as duas últimas foram incluídas aqui com a finalidade de encorajar futuros esforços experimentais para a descoberta de novos materiais. As propriedades de NTHs de tiofeno, furano e pirrol são tratadas com maior detalhe do que as de fosfol, uma vez que esta última apresenta dificuldades práticas de síntese (discutidas adiante). A nomenclatura utilizada para as 12 estruturas estudadas é uma combinação de sua configuração com o heteroátomo que ela apresenta. Por exemplo, as estruturas mostradas na Figura 17 para o caso do tiofeno são chamadas *S-syn*, *S-anti* e *S-syn-anti*. Para representar as demais, usam-se os prefixos N-, O- e P- para designar as estruturas derivadas de moléculas de pirrol, furano e fosfol, respectivamente.

Cuidados adicionais foram tomados em relação à otimização estrutural dos NTHs de pirrol e fosfol, uma vez que eles apresentam átomos de hidrogênio adicionais (nos grupos NH e

PH). Há diferentes mínimos locais de energia para a geometria das ligações N-H e P-H, nos quais os átomos de hidrogênio podem ocupar mais de uma posição. Todas as geometrias possíveis foram investigadas, e as propriedades relatadas são para as mais estáveis, a menos que seja afirmado o contrário.

4.1.2 Estabilidade Relativa

Com o intuito de analisar a estabilidade relativa das estruturas em estudo, foi calculada a sua Energia de Ligação (*Binding Energy*, E_b), definida como a diferença entre a energia do NTH relaxado (normalizada pelo número de moléculas equivalentes) e a de sua molécula de origem (ambos no vácuo e a 0K). Essa propriedade permite a comparação de estabilidade entre isômeros: quanto menor a E_b , mais estável a estrutura. Em particular, valores negativos de E_b indicam que o NTH é mais estável do que sua molécula precursora, o que é um importante indicativo de que a síntese do material é plausível do ponto de vista termodinâmico.

A Tabela 4 apresenta a E_b de todas as estruturas estudadas, bem como outras propriedades que serão discutidas a seguir. Na maioria dos casos, as configurações *anti* e *syn-anti* são mais estáveis que sua equivalente *syn*, uma vez que nesta última as interações de van der Waals entre heteroátomos subsequentes são predominantemente repulsivas. Seguindo essa lógica, as estruturas *anti* são as mais estáveis devido a uma menor repulsão, seguidas das *syn-anti*. Esse fenômeno é mais acentuado quando os heteroátomos são volumosos, como é o caso do enxofre. Por exemplo, a diferença de energia entre as configurações *anti* e *syn-anti* é maior para o caso do tiofeno do que do furano (24 versus 1 meV); e a estrutura S-*syn* chega a ter E_b positiva. O mesmo efeito é observado em NTHs derivados de benzeno e piridina, nos quais há repulsão entre os orbitais π das ligações duplas [9].

Essa maior estabilidade pode ser o motivo pelo qual as configurações *anti* e *syn-anti* estão mais presentes em amostras experimentais do que a *syn*, tanto em materiais derivados de moléculas de cinco membros quanto de seis [5, 7, 8, 10, 11]. No entanto, ao contrário dos casos de benzeno e piridina, as estruturas *anti* e *syn-anti* estudadas aqui são totalmente saturadas, possuindo E_b negativa devido à menor estabilização por ressonância em ciclos de cinco membros, o que facilita a sua saturação a pressões mais baixas [7, 8].

A única exceção encontrada nessa tendência é para os materiais derivados de pirrol: nesse caso, as configurações *syn* e *syn-anti* possuem estabilidade maior que a *anti*. Isso se deve à formação de ligações de hidrogênio entre grupos NH sucessivos, similarmente ao que acontece

NTH	E_b (eV)	Y_m (nN)	E_g^{GGA} (eV)	E_g^{HSE} (eV)
Square Polymer*	-0.465	103.2	4.25	–
N- <i>anti</i>	-0.285	61.6	3.65	4.76
N- <i>syn-anti</i>	-0.393	41.7	3.36	4.48
N- <i>syn</i>	-0.337	95.0	2.90	3.95
O- <i>anti</i>	-0.784	60.6	4.47	5.74
O- <i>syn-anti</i>	-0.783	38.4	4.16	5.46
O- <i>syn</i>	-0.625	90.2	3.32	4.61
S- <i>anti</i>	-0.461	62.2	3.61	4.84
S- <i>syn-anti</i>	-0.437	43.2	2.92	4.03
S- <i>syn</i>	+0.516	79.5	0.87	1.64
P- <i>anti</i>	-1.639	50.9	4.06	5.11
P- <i>syn-anti</i>	-1.592	34.5	3.68	4.65
P- <i>syn</i>	-0.935	110.8	1.52	2.20

TABELA 4 – Propriedades calculadas dos NTHs estudados. E_b é a Energia de Ligação, Y_m é o Módulo de Young 1D, E_g^{GGA} é o valor de *band gap* segundo cálculos de DFT GGA/PBE, e E_g^{HSE} é o *band gap* segundo DFT HSE. *As propriedades da estrutura de referência foram obtidos de Demingos e Muniz [9].

em NTHs de anilina [6]. No material *syn* esses grupos estão alinhados em um único lado da cadeia, e no *syn-anti* eles estão separados em pares. A distância entre um N e o H do próximo grupo NH é de 1.74 e 2.06 Å para as configurações *syn* e *syn-anti*, respectivamente, o que cai no intervalo característico de ligações de hidrogênio.

Com a finalidade de quantificar esse efeito, foram comparadas duas geometrias diferentes da estrutura N-*syn-anti*, movendo-se a posição dos átomos H de maneira a desfazer as ligações de hidrogênio (como mostrado na Figura 18). A estrutura da esquerda é a mais estável (todas as propriedades da configuração N-*syn-anti* foram calculadas a partir dela), com duas ligações de hidrogênio por célula unitária. A estrutura da direita, por sua vez, não apresenta nenhuma ligação de hidrogênio, de maneira que sua E_b é cerca de 0.2 eV maior. Esse resultado reforça que as ligações de hidrogênio são um importante fator de estabilização de NTHs derivados de pirrol, o que pode levar à síntese de estruturas que não se formariam a partir de outras moléculas aromáticas – isto é, uma maior quantidade de material na configuração *syn*.

A configuração *syn-anti* possui a E_b mais baixa entre os NTHs de pirrol, apesar de ter metade das ligações de hidrogênio, o que se deve à combinação desse efeito com a menor repulsão estérica entre os grupos funcionais anexados. De fato, essa repulsão tem um papel na desestabilização da estrutura N-*syn*: nesse material, a distância entre N e o próximo H é muito pequena, de apenas 1.74 Å, o que fica no limite inferior do intervalo característico de ligações de hidrogênio.

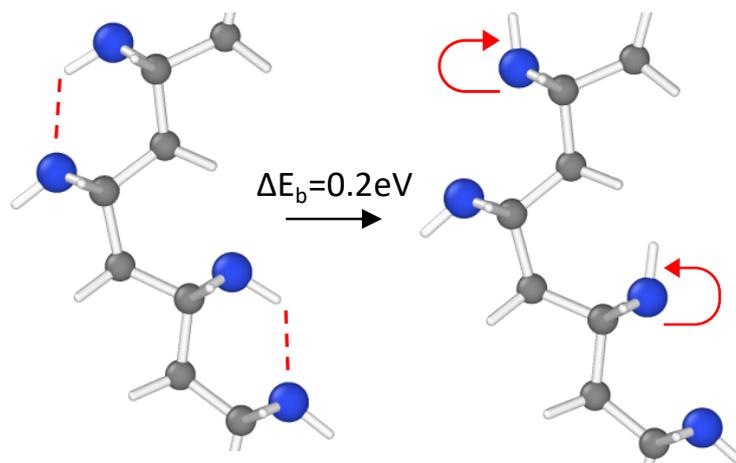


FIGURA 18 – Estrutura atômica do N-*syn-anti* em duas geometrias diferentes: com e sem ligações de hidrogênio. Reproduzido da referência [12] com permissão da Royal Society of Chemistry / PCCP Owner Societies.

Com a finalidade de verificar a estabilidade das estruturas em temperatura ambiente (300K), foram realizadas simulações de MD *ab initio*. Não houve quebra de ligações químicas, com estabilização da energia potencial depois de alguns picosegundos, como mostrado na Figura 19.

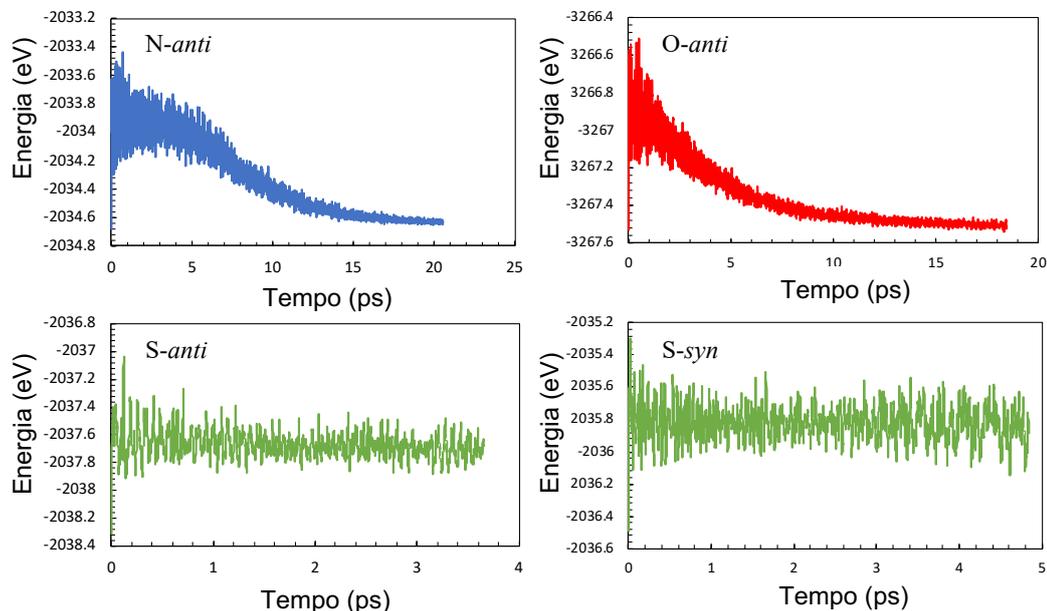


FIGURA 19 – Energia de NTHs representativos durante MD. Reproduzido da referência [12] (material suplementar) com permissão da Royal Society of Chemistry / PCCP Owner Societies.

4.1.3 Propriedades Eletrônicas

NTHs totalmente saturados derivados de benzeno são isolantes, com *band gap* superior a 3.5 eV segundo cálculos de DFT-GGA [4, 9] (esses valores ficam ca. 1.0 eV maiores quando são

feitos cálculos mais precisos empregando funcionais híbridos [47]). Isso se deve à hibridização sp^3 dos átomos dessas estruturas, que se assemelham a um filamento ultrafino de diamante. No entanto, amostras reais obtidas a partir do benzeno possuem carbono sp^2 , e é possível sintetizar NTHs a partir de moléculas contendo heteroátomos, como piridina e tiofeno. Trabalhos recentes mostram que a presença de ligações duplas C=C e C=N, bem como a presença de heteroátomos (principalmente átomos de nitrogênio com hibridização sp^3), reduz consideravelmente o *band gap* do material (de cerca de 4.00 eV para até 1.64 eV segundo DFT GGA/PBE) [9].

Com o objetivo de verificar se o mesmo ocorre em NTHs derivados de moléculas de cinco membros, foram calculadas suas propriedades eletrônicas. A Figura 20a mostra a estrutura de bandas dos materiais estudados, bem como sua densidade de estados (*Density of States*, DOS) total e projetada em orbitais. A Figura 20b mostra as propriedades do *Square Polymer* como referência, e a Tabela 4 apresenta os valores de *band gap*. Todos os NTHs estudados aqui apresentam *band gap* direto, e a presença de heteroátomos reduz o *band gap* em relação à estrutura derivada de benzeno. Essa redução se dá através da inserção de níveis ocupados de alta energia, diminuindo a distância entre o topo da banda de valência (*Valence Band Maximum*, VBM) e o fundo da banda de condução (*Conduction Band Minimum*, CBM), de maneira semelhante ao que ocorre em NTHs formados a partir de piridina e derivados de benzeno [9, 41]. Esse efeito fica visível a partir da PDOS, que mostra que essas bandas intermediárias estão ocupadas por elétrons provenientes dos heteroátomos. Esse processo é semelhante à dopagem do tipo n em semicondutores inorgânicos, apesar de não poder ser chamado dessa forma em função da grande quantidade de heteroátomos presentes. O efeito é mais pronunciado no caso do enxofre em função de seu maior volume, o que aumenta as interações entre os orbitais de átomos sucessivos (discutida adiante). Em particular, a estrutura S-*syn* apresenta o menor *band gap* relatado até então para NTHs de qualquer tipo: 0.87 eV (enquanto materiais derivados de moléculas de seis membros atingem um mínimo de 1.64 eV segundo DFT-GGA [9]). Isso sugere que NTHs derivados de tiofeno podem ter grande potencial como semicondutores 1D em aplicações práticas.

Comparando-se NTHs derivados da mesma molécula, a configuração *anti* possui o maior *band gap*, seguida da *syn-anti* e então da *syn*. Além disso, o *band gap* das estruturas *anti* não é tão dependente da espécie dos heteroátomos quanto as outras configurações: os *gaps* das estruturas *syn* variam de 0.87 a 3.32 eV, enquanto as *anti* variam apenas de 3.61 a 4.47 eV. Todos essas tendências ilustram a importância das interações entre grupos anexados nas propriedades

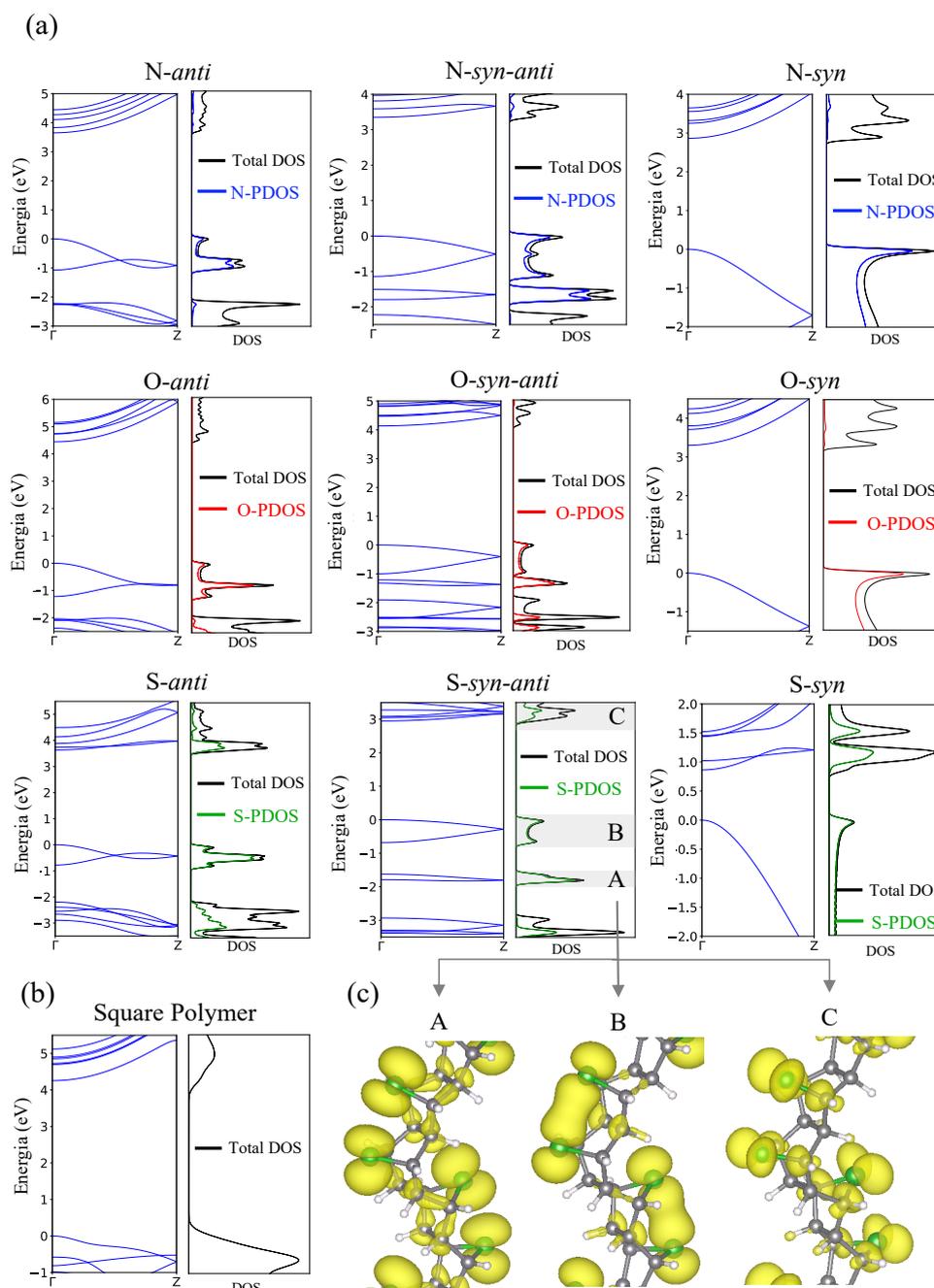


FIGURA 20 – Propriedades eletrônicas dos NTHs. Estrutura de bandas e DOS segundo DFT-GGA dos NTHs derivados de (a) moléculas de cinco membros e (b) Square Polymer. (c) ILDOS para três intervalos de energia da estrutura S-syn-anti. Reproduzido da referência [12] com permissão da Royal Society of Chemistry / PCCP Owner Societies.

eletrônicas dos NTHs. Um efeito parecido ocorre em NTHs parcialmente saturados de benzeno e piridina, nos quais os orbitais p_z das ligações π (C=C ou C=N) adjacentes interagem entre si. A Figura 20c ilustra esse efeito mostrando a densidade de estados local integrada (*Integrated Local Density of States*, ILDOS) do material S-syn-anti, representando espacialmente três intervalos de energia de sua estrutura de bandas: duas bandas ocupadas (A e B) introduzidas pelo enxofre,

e uma banda desocupada (C). Note que as regiões B e C incluem, respectivamente, a banda de valência e a banda de condução do material. Comparando-se os intervalos de energia A e B, fica claro que a banda relacionada ao orbital com maior interação entre grupos sucessivos (B) é aquela que tem um efeito mais pronunciado no aumento da VBM do material, reduzindo seu *band gap*. Outra consequência desse efeito pode ser vista na Figura 20a: estruturas *anti* e *syn-anti* apresentam bandas mais planas perto do nível de Fermi, uma vez que possuem menos interações entre grupos anexados, enquanto as estruturas *syn* apresentam uma banda de valência dispersa ao longo de um grande intervalo de energia (principalmente o material S-*syn* devido ao volume dos átomos de enxofre, como mencionado anteriormente).

De todas as estruturas estudadas, as derivadas do furano apresentam os maiores valores de *band gap*. Isso se deve ao maior grau de localização dos orbitais de seus heteroátomos, o que ocorre em função da grande diferença de eletronegatividade das ligações C-O. A Figura 21 mostra mapas de distribuição de carga em uma seção transversal aos materiais de configuração *syn*, confirmando que o oxigênio gera uma maior polarização das ligações químicas, seguido pelo nitrogênio do NTH derivado de pirrol. A presença desses grupos polares pode melhorar as interações dos NTHs com polímeros, otimizando sua aplicação como reforço em nanocompósitos, bem como sua solubilidade em solventes polares, facilitando seu pós-processamento.

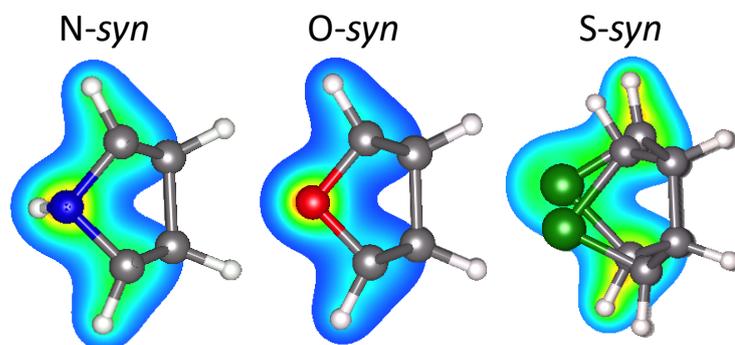


FIGURA 21 – Mapas de densidade de carga das estruturas *syn*. Reproduzido da referência [12] com permissão da Royal Society of Chemistry / PCCP Owner Societies.

A comparação dos resultados obtidos com as propriedades do *Square Polymer* permite analisar isoladamente os efeitos dos heteroátomos, uma vez que essa estrutura é totalmente saturada e formada somente por C e H. Outra comparação válida é com as estruturas *syn*, *anti* e *syn-anti* de benzeno, que, apesar de serem parcialmente saturadas, possuem uma estrutura ainda mais parecida com os NTHs estudados aqui. O *band gap* desses materiais está apresentado na Figura 22a. As faixas mostradas para NTHs de cada molécula partem da definição introduzida

na Seção 4.1.1, de que o número n de grupos anexados sucessivos (do mesmo lado da cadeia principal) pode ser utilizado para caracterizar essas estruturas. Nesse sentido, materiais com qualquer estrutura entre $n = 1$ (*anti*) e $n = \infty$ (*syn*) são teoricamente possíveis. Uma vez que o *band gap* da configuração com $n = 2$ (*syn-anti*) sempre cai entre os valores das anteriores, e em vista da discussão feita anteriormente sobre a importância das interações entre grupos anexados, é razoável pensar que outras estruturas ($n = 3, 4, 5, \dots$) terão *band gap* dentro das faixas mostradas. Essa análise mostra que as ligações duplas C=C em NTHs parcialmente saturados de benzeno têm um efeito mais pronunciado na redução do *band gap* do que os grupos O e NH, mas inferior ao S, o que reforça a possibilidade de materiais derivados de tiofeno terem grande potencial para aplicação como semicondutores.

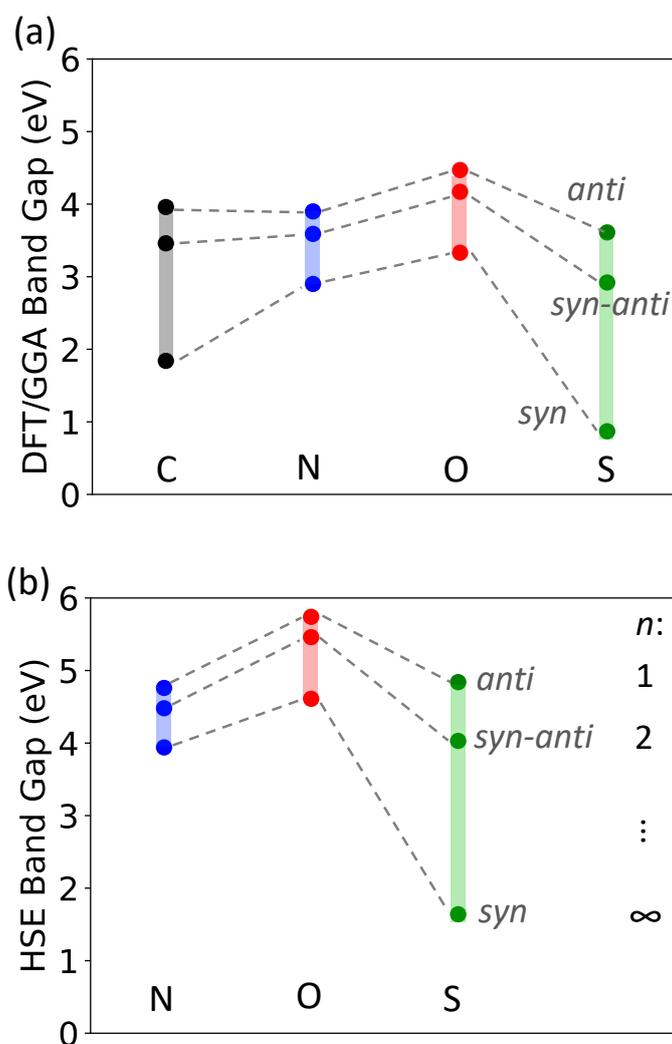


FIGURA 22 – Intervalos de *band gaps* de NTHs derivados de benzeno (C), pirrol (N), furano (O) e tiofeno (S), segundo (a) DFT-GGA, (b) DFT-HSE. Reproduzido da referência [12] com permissão da Royal Society of Chemistry / PCCP Owner Societies.

Uma vez que cálculos de DFT-GGA subestimam os valores de *band gap*, essa propriedade também foi computada utilizando-se o funcional híbrido HSE. Os resultados estão apresentados na Figura 22b, que utiliza a mesma ideia de faixas com base no número n , e também estão mostrados na Tabela 4. Todas as tendências são iguais às aquelas discutidas anteriormente para cálculos de DFT-GGA. A única alteração é um aumento nos valores absolutos de *gap*, que variam de 1.6 a 5.8 eV. Esses resultados mostram que NTHs derivados de moléculas de 5 membros tendem a apresentar comportamento isolante, e que o controle de sua estrutura pode levar a *band gaps* relativamente baixos, na faixa de semicondutores, o que abre um amplo leque de possíveis aplicações para esses materiais.

4.1.4 Propriedades Mecânicas

Estudos prévios mostram que NTHs derivados de moléculas de seis membros possuem excelentes propriedades mecânicas [1, 2]. Além disso, a presença de heteroátomos não altera significativamente a resistência desses materiais, especialmente quando não estão localizados na cadeia principal, mas ligados na sua superfície como apêndices [9, 41]. Com o objetivo de verificar se o mesmo ocorre em NTHs derivados de moléculas de cinco membros, foram simuladas curvas de tensão-deformação (Figura 23a), a partir das quais foi calculado o Módulo de Young 1D de cada material (mostrados na Tabela 4), seguindo a metodologia descrita na Seção 3.1.

A rigidez dos materiais (medida pelo Módulo de Young) varia mais fortemente com a configuração da cadeia (*syn*, *anti* ou *syn-anti*) do que com o heteroátomo presente, o que fica evidenciado pela parte inicial das curvas de tensão-deformação. Os valores obtidos também são parecidos com os das estruturas equivalentes derivadas de benzeno e piridina (por exemplo, as estruturas *anti*, *syn-anti* e *syn* de benzeno possuem Módulo de Young de 60.9, 40.0 e 82.2 nN, respectivamente [9]). Esses resultados confirmam que a rigidez de NTHs é basicamente definida por sua cadeia principal, sendo fracamente afetada por grupos funcionais anexados. Além disso, a maior variação do Módulo de Young causada pela espécie dos heteroátomos ocorre entre as configurações *syn*: enquanto as variações das outras configurações são inferiores a 5.0 nN, o Módulo de Young do S-*syn* é até 15.0 nN inferior às outras estruturas *syn*. Isso ocorre porque essa configuração apresenta maior repulsão entre os grupos anexados, como discutido anteriormente.

Todas as estruturas estudadas apresentaram comportamento frágil, de maneira semelhante a outros NTHs [1, 2, 9]. As configurações *anti* e *syn-anti* suportam deformações maiores

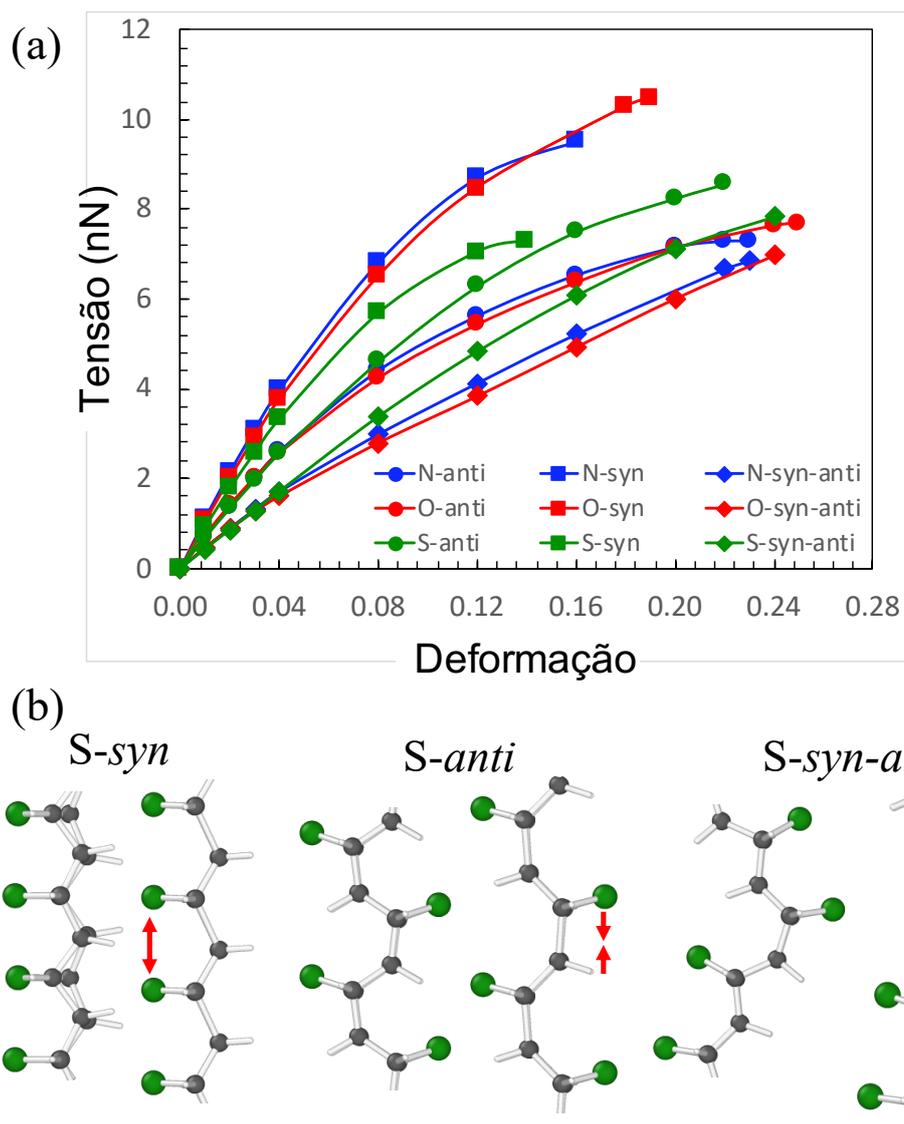


FIGURA 23 – Propriedades mecânicas dos NTHs. (a) Curvas de tensão-deformação simuladas por DFT (linhas servem como guia visual). (b) Estrutura atômica dos NTHs derivados de tiofeno relaxados (esquerda de cada par) e deformados (direita); as setas vermelhas representam o movimento dos átomos anexados devido ao estiramento do material. Reproduzido da referência [12] com permissão da Royal Society of Chemistry / PCCP Owner Societies.

antes da fratura (até 0.25), o que é devido à sua estrutura corrugada, que permite deformação angular das ligações antes de seu estiramento, como mostrado na Figura 23b. O mesmo efeito foi observado em estruturas semelhantes derivadas de benzeno e piridina, com valores semelhantes de deformação máxima suportada [9]. Além disso, as estruturas *syn* tendem a ser mais resistentes e rígidas do que outras configurações com o mesmo heteroátomo. Isso é devido a um melhor alinhamento das ligações C-C à direção de aplicação da tensão (eixo z da célula unitária). Em estruturas *syn*, o estiramento das ligações é o mecanismo principal de deformação mesmo a baixos valores de tensão – o contrário do que acontece nas outras configurações. A tensão de ruptura

1D dos NTHs estudados aqui estão na faixa de 6.0-11.0 nN, similarmente a estruturas derivadas de moléculas de seis membros [9]. Esses resultados reforçam a ideia de que as propriedades mecânicas de NTHs têm pouca dependência de grupos anexados, sugerindo que NTHs derivados de moléculas de cinco átomos também podem ser usados em aplicações que exigem resistência mecânica, como no reforço de nanocompósitos.

No entanto, NTHs derivados do tiofeno apresentaram comportamento diferente dos demais NTHs, como mostrado na Figura 23a. Enquanto as estruturas de furano e pirrol possuem curvas de tensão-deformação quase iguais, os NTHs com enxofre se desviam em altos valores de deformação (acima de ca. 0.08), mantendo-se abaixo ou acima das outras curvas de mesma configuração. Esse efeito é atribuído ao maior volume dos átomos de enxofre e à consequente repulsão entre eles. Nas estruturas *syn*, a deformação axial afasta os heteroátomos (setas vermelhas da Figura 23b), de maneira que a repulsão entre eles atua como uma força interna adicional que facilita o estiramento. Como consequência, é necessária uma força externa menor para deformar a estrutura; logo, o material S-*syn* é menos resistente que seus semelhantes. Por outro lado, nas configurações *anti* e *syn-anti* a deformação do material leva a uma aproximação dos heteroátomos (ou do heteroátomo e átomos de hidrogênio). Nesse caso, a repulsão do enxofre atua como uma resistência adicional ao estiramento; logo, os materiais S-*anti* e S-*syn-anti* são mais resistentes que seus semelhantes (Figura 23a).

4.1.5 Nanofilamentos de carbono derivados de fosfol

A síntese de NTHs a partir de fosfol apresenta uma dificuldade adicional: a instabilidade da molécula precursora. Em condições normais, o fosfol forma dímeros através de transferência de próton [126], dando origem a moléculas maiores que não poderiam ser utilizadas na formação das configurações *syn*, *anti* e *syn-anti* (alguns desses dímeros possuem, por exemplo, ligações P-P). No entanto, com a finalidade de apresentar um conjunto mais completo de resultados sobre NTHs derivados de moléculas de cinco membros, as mesmas três configurações foram estudadas aqui, e suas propriedades são brevemente discutidas a seguir. Esses resultados podem ser utilizados no futuro como uma referência para análises mais aprofundadas de NTHs derivados de fosfol. (Pirrol também tem tendência a oligomerizar, mas em uma escala de tempo muito maior, o que permite sua purificação para síntese de NTHs.)

A Figura 24 mostra as estruturas *syn*, *anti* e *syn-anti* derivadas de fosfol, bem como suas propriedades eletrônicas. As estruturas propostas são equivalentes às do pirrol, mas com os

átomos de N substituídos por P. Além disso, a estrutura *P-syn* apresenta um desvio em relação à configuração ideal: durante a otimização estrutural, os átomos H dos grupos PH se moveram para o plano das ligações C-P-C. Enquanto nos NTHs de pirrol os grupos NH formam ligações de hidrogênio, no caso do fosfol as ligações P-H são grandes demais para acomodar uma geometria não-planar na configuração *syn*. No entanto, uma geometria tetraédrica é observada nos materiais *P-anti* e *P-syn-anti*, como esperado.

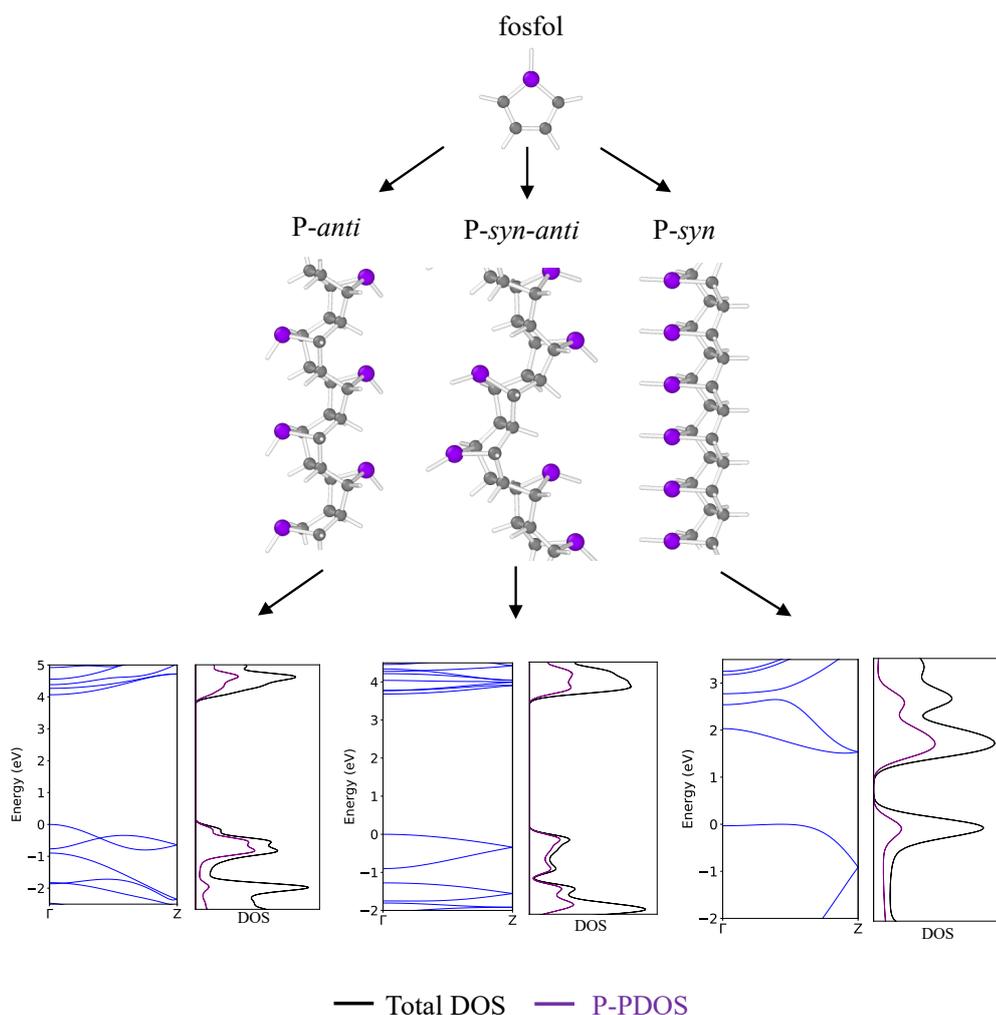


FIGURA 24 – Estruturas e propriedades eletrônicas dos NTHs derivados de fosfol (estrutura de bandas e DOS). Reproduzido da referência [12] (material suplementar) com permissão da Royal Society of Chemistry / PCCP Owner Societies.

Adicionalmente às estruturas de bandas e densidades de estados eletrônicas mostradas na Figura 24, a Tabela 4 apresenta as propriedades calculadas para NTHs de fosfol: energia de ligação E_b , *band gap* segundo DFT-GGA e HSE, e Módulo de Young 1D. Todos os valores seguem as tendências discutidas anteriormente para as outras moléculas. A única característica digna de nota é a baixa E_b dos NTHs de fosfol, quando comparados aos demais. Isso se deve

não apenas à estabilidade dessas estruturas, mas também à instabilidade da molécula precursora (tomada como referência para cálculo da E_b).

4.2 DESENVOLVIMENTO DE UM ALGORITMO GENÉTICO PARA EXPLORAÇÃO DE ESTRUTURAS COMPLEXAS DE NANOFILAMENTOS DE CARBONO

4.2.1 Procura por nanofilamentos de benzeno com alta estabilidade

Com o objetivo de demonstrar que o algoritmo genético implementado pode ser utilizado para buscar e otimizar estruturas complexas de NTHs, foi realizada uma exploração extensiva (diversos processos evolutivos independentes) em NTHs derivados de benzeno, segundo a metodologia descrita na Seção 3.2.7. Foram utilizadas estruturas não-periódicas de diferentes tamanhos, formadas por 4, 8 e 16 moléculas. A Figura 25 mostra alguns exemplos dos NTHs gerados. A relaxação estrutural foi feita através da minimização da energia via método do gradiente conjugado (método padrão do comando *minimize* do LAMMPS), que permite uma otimização local. Por esse motivo, as estruturas mostradas na Figura 25 (e todas as outras simuladas) não se afastaram muito da conformação inicial, isto é, mantendo uma forma estendida/reta. Uma vez que NTHs são sintetizados em estado sólido, dando origem a uma fase cristalina, é razoável pensar que as conformações obtidas através do método usado aqui formam um bom modelo do material real, pelo menos quanto à forma estendida/reta dos NTHs. Uma otimização adicional poderia ser feita através de simulações MD a alta temperatura, que forneceriam estruturas em outros mínimos de energia, possivelmente com maior grau de empacotamento no caso de NTHs longos, de maneira semelhante ao que ocorre com polímeros.

A Figura 26 mostra dados de uma única evolução: a energia dos três candidatos mais estáveis ao longo das 20 gerações, bem como a média de energia de todos os candidatos desse processo evolutivo. O algoritmo genético minimizou a energia dos NTHs de maneira aproximadamente exponencial, como esperado. A utilização da singularidade no cálculo da probabilidade de seleção fica evidenciada na quarta geração, na qual o candidato mais estável aumenta de energia (tem uma diminuição de estabilidade). Esse evento ilustra a capacidade do algoritmo de fugir de mínimos locais, procurando regiões ainda mais promissoras. De fato, a maior queda na energia do candidato mais estável ocorre logo em seguida desse aumento, da quinta para a sexta geração, sugerindo que o algoritmo encontrou um novo mínimo. A maior variação na energia dos candidatos 2 e 3 ocorre na geração seguinte: muitos NTHs da sétima geração foram gerados a partir de mutações do candidato mais estável da sexta.

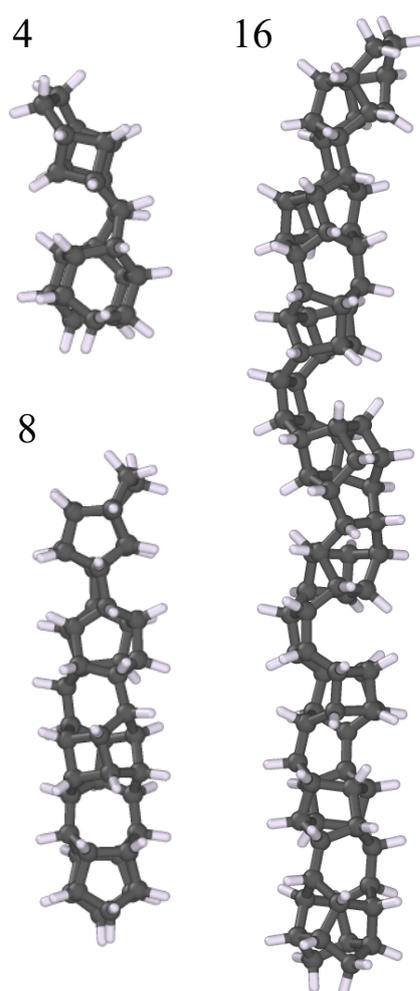


FIGURA 25 – Estruturas representativas de NTHs de benzeno com 4, 8 e 16 moléculas.

Em sua procura por NTHs estáveis, o algoritmo genético beneficia certas características estruturais em detrimento de outras (como ilustrado na Figura 25, onde são mostrados três NTHs da última geração de seus processos evolutivos). Por exemplo, ciclos de seis e cinco membros aparecem com maior frequência nos materiais, enquanto ciclos de quatro membros se tornam mais raros em função da alta energia causada pela geometria pouco favorável de seus átomos. Por exemplo, estruturas parecidas ao NTH ideal *Polymer I* aparecem com frequência em materiais otimizados, como mostrado no NTH de 16 moléculas da Figura 25, uma vez que essa estrutura apresenta alta estabilidade devido a seus múltiplos ciclos de cinco membros [4]. O NTH de 16 moléculas também apresenta estruturas locais em forma de C, parecidas com o que se observa em alguns NTHs ideais de classe II [4], e que dão ao material alta flexibilidade e resistência à tração [2].

Apesar da tendência descrita acima, ciclos de quatro membros não desaparecem completamente, como mostrado no NTH de 4 moléculas da Figura 25. De fato, esses ciclos "qua-

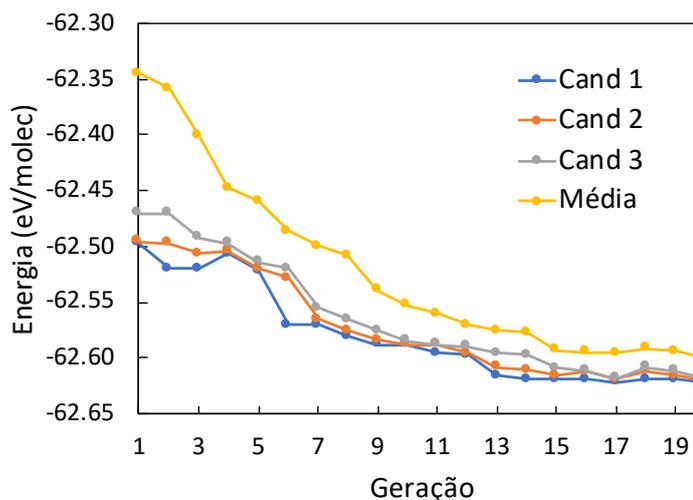


FIGURA 26 – Energia média e dos três candidatos mais estáveis em um processo evolutivo com NTHs de benzeno com 8 moléculas.

drados" aparecem em uma das estruturas ideais de maior estabilidade, o *Square Polymer* [4], que também é observado em amostras reais derivadas de benzeno [10, 11]. Essa estrutura, que possui um único ciclo de quatro membros, aparece com frequência em materiais otimizados pelo algoritmo genético, como na parte superior do NTH de 8 moléculas da Figura 25. O que é de fato removido dos NTHs pelo algoritmo genético é a ocorrência de múltiplos ciclos de quatro membros lado a lado. Em estruturas geradas aleatoriamente (na primeira geração, por exemplo), é comum que muitas ligações "diretas" sejam produzidas – isto é, átomos que se ligam à mesma posição na molécula superior. Duas ligações "diretas" consecutivas dão origem a um ciclo de quatro membros; três ligações consecutivas dão origem a dois ciclos adjacentes. Em NTHs de classe II, até quatro ligações podem ser feitas entre duas moléculas, o que pode gerar três ciclos "quadrados" lado a lado. Essas estruturas são comuns em genótipos produzidos aleatoriamente, mas não aparecem em NTHs otimizados pelo algoritmo, uma vez que possuem alta energia.

Uma vez que algoritmos genéticos são métodos metaheurísticos de otimização, em geral é necessário realizar múltiplas evoluções independentes para averiguar sua performance [99]. Em função disso, foi feita uma investigação mais extensiva do design de NTHs estáveis, segundo a metodologia descrita na Seção 3.2.7. Para cada tamanho de NTH, foram realizadas 100 evoluções independentes, cada uma com 400 candidatos. Para comparação, também foram simulados ($100 \times 400 =$) 40000 candidatos aleatórios de cada tamanho. A Figura 27 mostra as distribuições de energia obtidas. A comparação dos histogramas mostra que, em todos os casos, o algoritmo genético gera candidatos mais estáveis do que uma procura aleatória.

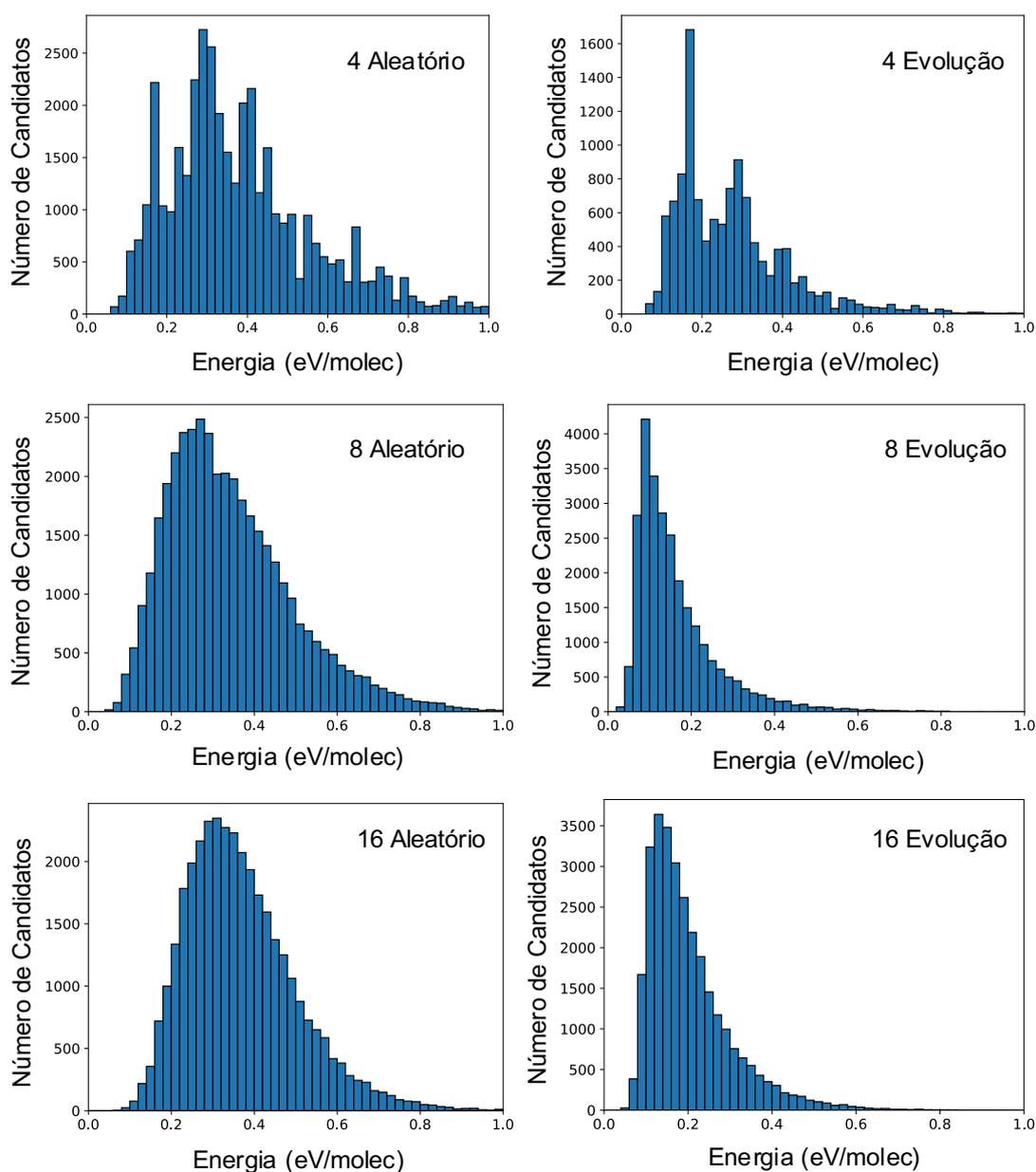


FIGURA 27 – Distribuição de energia relativa dos candidatos simulados via processo evolutivo e procura aleatória, com 4, 8 e 16 moléculas.

A energia destes NTHs não pode ser comparada diretamente com a das estruturas ideais apresentadas anteriormente (Figura 16, Seção 3.2.7), uma vez que aquelas são periódicas, enquanto estas possuem hidrogênios nas pontas. No entanto, as variações de energia são parecidas: da ordem de 0.1 eV por molécula. Além disso, apesar de o algoritmo genético procurar favorecer estruturas de menor energia (isto é, que se encontram mais à esquerda nos histogramas), a maioria dos candidatos simulados apresenta estabilidade mediana (evidenciado pelo máximo dos histogramas, que ocorre em energias superiores às dos candidatos mais estáveis). Isso ocorre porque, entre as estruturas possíveis, há mais NTHs com energias intermediárias do

que baixas (como ilustrado pelos histogramas gerados via procura aleatória). Uma vez que os histogramas mostrados só contabilizam cada candidato uma vez, mesmo que ele tenha sido replicado repetidamente durante a evolução, esses poucos NTHs de menor energia representam uma fração menor da população do que as estruturas de estabilidade mediana. Quanto às características estruturais, os candidatos com estabilidade média seguem as mesmas tendências dos otimizados, mas em menor escala: possuem menos ciclos de quatro membros, e mais de cinco e de seis.

A Figura 27 mostra algumas outras características do algoritmo genético. Pode ser visto, por exemplo, que a exploração de NTHs pequenos (4 moléculas) tem performance inferior à de estruturas maiores. Isso se deve ao menor número de graus de liberdade de NTHs curtos, o que impossibilita atingir materiais com estruturas mais complexas. Outro motivo é o maior grau de aleatorização que o algoritmo gera em NTHs pequenos: com apenas 4 moléculas, uma mutação altera um quarto do código genético (inclusive em cruzamentos simples), o que dificulta a otimização. Além disso, estruturas menores possuem uma distribuição de energia mais irregular, o que fica evidente ao compararmos os três histogramas dos materiais gerados aleatoriamente. Os NTHs de 16 moléculas formam uma distribuição suave, enquanto estruturas menores dão origem a imperfeições, como o "dente" em torno de 0.3 eV/molec nas estruturas de 8 moléculas e os diversos máximos e mínimos das estruturas de 4 moléculas. Essas irregularidades são consequência da natureza discreta do genótipo implementado – isto é, do fato de que há um número finito de possíveis padrões de ligações entre as moléculas –, o que dá um caráter semi-discreto às contribuições de cada gene para a energia total do material. Quanto maior o NTH, menos importa essa discretização, de maneira que a energia se torna uma função mais aproximadamente contínua do genótipo conforme o número de moléculas aumenta.

Uma vez que as evoluções foram realizadas com replicação (4 dos 20 candidatos de cada geração), os histogramas à direita da Figura 27 não possuem de fato 40000 NTHs diferentes. Em particular, a evolução com 4 moléculas possui um número visivelmente inferior de candidatos simulados (isto é, diferentes uns dos outros). Isso se deve ao número relativamente menor de genótipos possíveis neste caso, o que fez com que o algoritmo gerasse o mesmo NTH diversas vezes a partir da mutação de estruturas parecidas. Esse efeito, somado à irregularidade da distribuição de energia mencionada acima, sugere que o algoritmo genético não é apropriado para o estudo de NTHs pequenos. No entanto, uma vez que a implementação foi feita com o objetivo de explorar estruturas complexas, essa limitação não representa um problema. De

fato, NTHs formados por 4 moléculas podem ser explorados sem um método metaheurístico de otimização, e as propriedades da maioria deles já foram reportadas [2, 4, 9]. O uso deste algoritmo genético fica, então, reservado a NTHs maiores (com 8 moléculas ou mais).

4.2.2 Procura por nanofilamentos de benzeno com alta rigidez

Com a finalidade de demonstrar que o algoritmo genético implementado pode ser usado para otimizar diferentes propriedades de NTHs, foi realizada uma procura por estruturas com alto Módulo de Young (alta rigidez), segundo a metodologia descrita na Seção 3.2.7, com NTHs periódicos de 8 moléculas (a periodicidade é necessária no método utilizado para simular um teste de deformação uniaxial e medir o Módulo de Young). A Figura 28 mostra o Módulo de Young 1D do candidato mais rígido ao longo do processo evolutivo. Os valores obtidos vão de ca. 69 a 85 nN. Para comparação, o NTH ideal mais rígido é o *Tube(3,0)*, com Módulo de Young de 136 nN segundo ReaxFC [2]; no entanto, esse é um NTH se classe III, enquanto os materiais simulados aqui são de classe II, que tendem a possuir maior flexibilidade em função de suas múltiplas seções transversais com apenas duas ligações. Por exemplo, cálculos via ReaxFC mostraram que NTHs ideais de classe II (entre eles o *Polymer I*) possuem Módulo de Young na faixa de 83-88 nN [2], muito semelhante aos valores otimizados obtidos aqui.

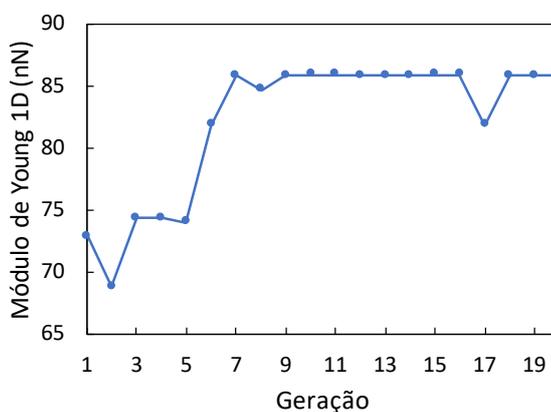


FIGURA 28 – Módulo de Young 1D do candidato mais rígido durante processo evolutivo para otimização da rigidez.

Na procura por NTHs rígidos, o algoritmo genético beneficiou ciclos de quatro membros, ao contrário do que foi observado anteriormente na otimização da estabilidade. Isso ocorre porque neste caso a alta energia dessas estruturas não é computada no cálculo da aptidão dos candidatos; em vez disso, a baixa flexibilidade desses ciclos causa um alto Módulo de Young, de maneira que eles são favorecidos. Em explorações futuras por NTHs para aplicações mecânicas práticas, pode

ser benéfico incluir dois termos no cálculo da aptidão, um dos quais quantifique a estabilidade do candidato, uma vez que materiais instáveis têm utilidade limitada, ainda que suas outras propriedades sejam excelentes.

Um algoritmo genético não necessariamente terá uma boa performance na otimização de qualquer propriedade. Como discutido na Seção 2.3, é necessário que pequenas alterações no genótipo produzam pequenas variações na aptidão, o que em geral ocorre quando as partes do código genético contribuem com parcelas (relativamente) independentes do valor de aptidão. Esse é o caso da energia do sistema, uma vez que esta depende da geometria local dos átomos. No caso da rigidez, foi demonstrado que NTHs compostos – isto é, produzidos pela combinação de estruturas idealizadas – podem ser modelados como um conjunto de molas em série [1], de maneira que sua elasticidade E pode ser aproximada por uma expressão na forma da Equação 4.1, onde E_i é a elasticidade de suas partes. Isso sugere que algoritmos genéticos constituem um método apropriado para a otimização da rigidez de NTHs. Ainda assim, é necessário um estudo mais minucioso dessa questão, semelhante ao que foi feito anteriormente para o caso da energia.

$$\frac{1}{E} = \sum_i \frac{1}{E_i} \quad (4.1)$$

4.2.3 Simulação de nanofilamentos de outras moléculas aromáticas

Visando demonstrar que o algoritmo implementado é capaz de gerar NTHs derivados de diferentes moléculas aromáticas, foram realizadas buscas por estruturas de alta estabilidade (baixa energia) para diferentes precursores. A metodologia utilizada no processo evolutivo foi aquela descrita na Seção 3.2.7 para NTHs de 8 moléculas, com a exceção de que neste caso as moléculas tiveram liberdade para rotacionar (uma vez que elas não possuem a simetria rotacional do benzeno). A Figura 29 mostra NTHs representativos de diferentes moléculas aromáticas, incluindo de cinco membros (pirrol, furano e tiofeno) e de seis membros (anilina e piridina). Os precursores de cinco membros fazem duas ligações com cada molécula vizinha, enquanto os de seis membros são de classe II, isto é, fazem quatro e duas ligações com seus vizinhos. Em função disso, os NTHs derivados de anilina e piridina tendem a ter uma estrutura mais complexa do que os das outras três moléculas.

Como mencionado anteriormente, as coordenadas atômicas dos NTHs foram otimizadas localmente, de maneira que as estruturas não têm liberdade para flexionar muito. Ainda assim, grandes átomos (como S) e grupos funcionais (como NH_2) causam flexões maiores, como

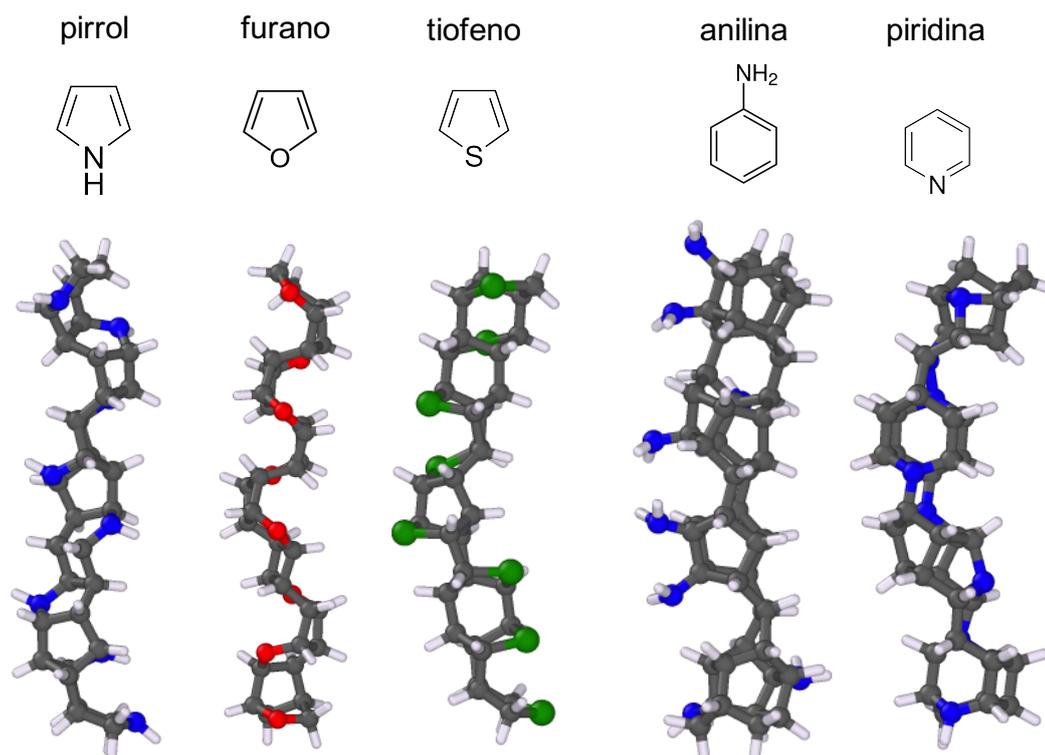


FIGURA 29 – Estruturas representativas de NTHs derivados de outras moléculas aromáticas.

demonstrado pelos NTHs de tiofeno e anilina. Isso sugere que NTHs contendo heteroátomos podem apresentar diferentes conformações quando colocado em solução (aquosa ou não), como anéis e espirais [127]. Uma vez que o algoritmo genético apresentado aqui pode modelar NTHs não-periódicos, ele pode ser utilizado para explorar essas conformações.

A Figura 30 mostra a energia relativa do candidato mais estável durante o processo evolutivo para diferentes moléculas aromáticas. O algoritmo evolutivo foi capaz de melhorar a estabilidade em todos os casos. As variações de energia não podem ser comparadas entre moléculas, uma vez que cada sistema possui um valor característico (o zero é diferente em cada caso, e os processos evolutivos foram apresentados em conjunto por simplicidade).

Em cada caso estudado, o algoritmo genético privilegiou estruturas que causam maior estabilidade. Por exemplo, ligações nitrogênio-nitrogênio são evitadas em NTHs de piridina devido à sua instabilidade [9], e grupos anexados tendem a aparecer de maneira alternada em NTHs derivados de tiofeno e furano para minimizar sua repulsão estérica [12]. Em estruturas de anilina, por outro lado, os grupos NH_2 aparecem em posições relativamente próximas, uma vez que podem formar ligações de hidrogênio entre si [6]. Em NTHs derivados de moléculas de cinco membros, estruturas formadas por cicloadições [4 + 2] aparecem com frequência em

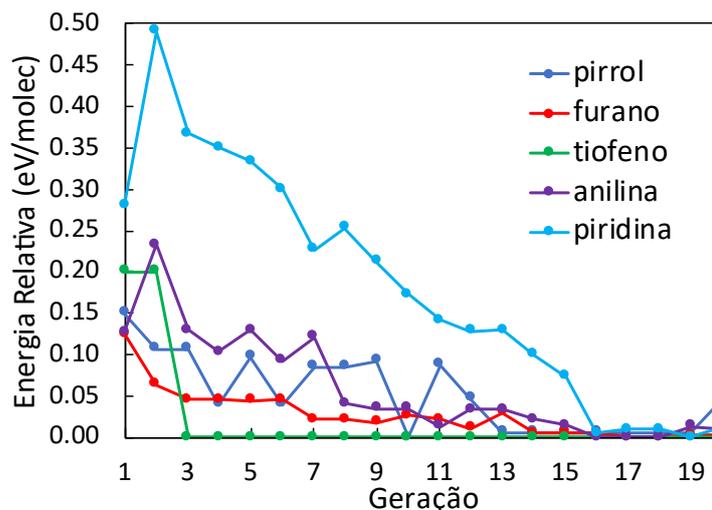


FIGURA 30 – Energia do candidato mais estável em evoluções para diferentes precursores. O zero de energia foi tomado como o candidato mais estável obtido em cada caso.

materiais otimizados. Tomando como exemplo o material de pirrol da Figura 29, a molécula mais inferior se liga à próxima pelos dois carbonos mais distantes do grupo NH, o que caracteriza as configurações *syn*, *anti* e *syn-anti*, discutidas na Seção 4.1.

Mais uma vez, é necessário um estudo mais minucioso para averiguar a performance do algoritmo genético ao lidar com diferentes moléculas precursoras, sendo possível que nesses casos outro campo de força produza resultados mais confiáveis que o ReaxFC. Os dados apresentados aqui têm como objetivo somente demonstrar essa funcionalidade. Ainda assim, métodos metaheurísticos são menos necessários na investigação de NTHs derivados de moléculas de cinco membros, uma vez que eles possuem estruturas mais organizadas [8]. Entretanto, uma vez que a síntese de NTHs é controlada cineticamente [3, 4, 19], novas estruturas desses NTHs podem ser desejáveis no futuro, de maneira que um método computacional de otimização será necessário para guiar o design desses materiais. O algoritmo genético apresentado aqui pode ser utilizado para esse fim.

5 CONCLUSÕES

A investigação computacional de nanofilamentos de carbono (NTHs) proposta nesta dissertação foi realizada em duas partes. Na primeira, cálculos de primeiros princípios elucidaram algumas propriedades de NTHs derivados de moléculas aromáticas heterocíclicas de cinco membros. Na segunda, foi desenvolvido um algoritmo genético para a otimização de estruturas complexas de NTHs derivados de diferentes compostos aromáticos.

Cálculos DFT foram empregados na investigação de propriedades estruturais, eletrônicas e mecânicas de NTHs derivados de moléculas de furano, tiofeno, pirrol e fosfol. Os resultados mostraram que esses materiais possuem alta estabilidade: a maioria das estruturas são mais estáveis do que suas moléculas de origem, inclusive algumas configurações *syn*, o que não ocorre para NTHs investigados anteriormente na literatura. Isso abre a possibilidade da síntese de materiais com estruturas ainda não obtidas em laboratório, em particular no caso de NTHs derivados de pirrol, que são estabilizados pela formação de ligações de hidrogênio internas. Também observou-se que a presença de heteroátomos introduz níveis ocupados de alta energia na estrutura eletrônica do material, reduzindo seu *band gap* até valores inferiores àqueles relatados na literatura para NTHs de outras moléculas. Em alguns casos, os materiais tornam-se semicondutores, principalmente no caso de estruturas derivadas de tiofeno, o que expande o leque de possíveis aplicações para NTHs. Por fim, verificou-se que NTHs derivados de anéis de cinco membros possuem propriedades mecânicas semelhantes às das estruturas obtidas a partir de benzeno e piridina, o que garante sua aplicabilidade nas áreas já sugeridas e computacionalmente testadas para esses materiais, como em nanofibras e reforço de compósitos. Além disso, a presença de heteroátomos volumosos (como enxofre) altera levemente a resistência à tração de NTHs, aumentando-a (nas configurações *anti* e *syn-anti*) ou diminuindo-a (na *syn*), dependendo do movimento relativo que os grupos anexados fazem durante o estiramento do material.

Os dados reportados nesta parte da dissertação também foram reportados no artigo *First-principles study of carbon nanothreads derived from five-membered heterocyclic rings: thiphene, furan and pyrrole*, publicado na *Physical Chemistry Chemical Physics* em 2021 (*Phys. Chem. Chem. Phys.*, 23, 2055). Espera-se que esses resultados encorajem mais esforços experimentais no sentido de sintetizar novos nanomateriais, como NTHs derivados de pirrol. Trabalhos futuros

poderão explorar o transporte eletrônico em NTHs, inclusive com a presença de diferentes heteroátomos.

O algoritmo genético proposto foi implementado com sucesso. Ele permite a variação de diversos parâmetros de otimização, como a molécula precursora, o número de moléculas por estrutura, o número de estruturas (candidatos) por geração, o número de gerações, a propriedade a ser otimizada (aptidão), e o método de simulação computacional empregado para calculá-la, entre outros parâmetros. A performance do algoritmo foi testada em uma busca extensiva por NTHs de alta estabilidade; quando comparado a uma procura aleatória, ele se mostrou capaz de produzir estruturas mais estáveis (com maior aptidão). Além disso, a capacidade do algoritmo de otimizar diferentes propriedades foi demonstrada através do design de NTHs de alta rigidez (alto Módulo de Young). O algoritmo genético apresentado aqui pode ser utilizado para elucidar a estrutura de amostras reais de NTHs (comparando-se o material simulado ao real e minimizando a diferença), bem como para otimizar alguma propriedade de interesse (cuja estrutura pode ser produzida na prática alterando-se as condições de síntese).

O algoritmo genético foi implementado como parte do pacote ChemStruct, cuja primeira versão está disponibilizada em Python desde 2019 (tanto em PyPI quanto pelo link <https://github.com/pdemingos/chemstruct>). A próxima versão do ChemStruct incluirá o algoritmo genético revisado, que poderá ser utilizado livremente para o design e otimização de NTHs. Para a apresentação do algoritmo em artigo, uma investigação mais completa de seus parâmetros será realizada, bem como das propriedades investigadas e dos métodos utilizados para computá-las (incluindo DFT). Também poderá ser feita uma comparação do algoritmo genético com outros métodos de otimização, como métodos locais (utilizando, por exemplo, uma equação empírica) ou Monte Carlo (realizando "mutações" unitárias nos NTHs e estimando a sua energia). Expansões do algoritmo poderão ser realizadas para a simulação de NTHs parcialmente saturados (com ligações duplas) e em estado sólido (cristais hexagonais).

REFERÊNCIAS

- [1] Ruth E Roman, Kenny Kwan e Steven W Cranford. “Mechanical properties and defect sensitivity of diamond nanothreads”. Em: *Nano letters* 15.3 (2015), pp. 1585–1590.
- [2] Julian FRV Silveira e Andre R Muniz. “First-principles calculation of the mechanical properties of diamond nanothreads”. Em: *Carbon* 113 (2017), pp. 260–265.
- [3] Thomas C Fitzgibbons et al. “Benzene-derived carbon nanothreads”. Em: *Nature materials* 14.1 (2015), p. 43.
- [4] En-shi Xu, Paul E Lammert e Vincent H Crespi. “Systematic enumeration of sp³ nanothreads”. Em: *Nano letters* 15.8 (2015), pp. 5124–5130.
- [5] Xiang Li et al. “Carbon Nitride Nanowire Crystals Derived from Pyridine”. Em: *Journal of the American Chemical Society* 140.15 (2018), pp. 4969–4972.
- [6] Marcelo M Nobrega et al. “One-dimensional diamondoid polyaniline-like nanowires from compressed crystal aniline”. Em: *Chemical science* 9.1 (2018), pp. 254–260.
- [7] Steven Huss et al. “Scalable Synthesis of Crystalline One-Dimensional Carbon Nanowires through Modest-Pressure Polymerization of Furan”. Em: *ACS nano* 15.3 (2021), pp. 4134–4143.
- [8] Arani Biswas et al. “Evidence for Orientational Order in Nanowires Derived from Thiophene”. Em: *The journal of physical chemistry letters* 10.22 (2019), pp. 7164–7171.
- [9] Pedro G Demingos e Andre R Muniz. “Electronic and Mechanical Properties of Partially Saturated Carbon and Carbon Nitride Nanowires”. Em: *The Journal of Physical Chemistry C* 123.6 (2019), pp. 3886–3891.
- [10] Pu Duan et al. “The Chemical Structure of Carbon Nanowires Analyzed by Advanced Solid-State NMR”. Em: *Journal of the American Chemical Society* (2018).
- [11] Tao Wang et al. “Constraining carbon nanowire structures by experimental and calculated nuclear magnetic resonance spectra”. Em: *Nano letters* 18.8 (2018), pp. 4934–4942.

- [12] Pedro G Demingos, Naira M Balzaretti e André R Muniz. “First-principles study of carbon nanothreads derived from five-membered heterocyclic rings: thiophene, furan and pyrrole”. Em: *Physical Chemistry Chemical Physics* 23.3 (2021), pp. 2055–2062.
- [13] Xiang Li et al. “Mechanochemical synthesis of carbon nanothread single crystals”. Em: *Journal of the American Chemical Society* 139.45 (2017), pp. 16343–16349.
- [14] Weizhao Cai et al. “Deuterium isotope effects in polymerization of benzene under pressure”. Em: *The Journal of Physical Chemistry Letters* 8.8 (2017), pp. 1856–1864.
- [15] Ayako Shinozaki et al. “Pressure-induced irreversible amorphization of naphthalene and nitrogen-containing heteroaromatic compounds at room temperature”. Em: *Chemical Physics Letters* 739 (2020), p. 136921.
- [16] Dragan Stojkovic, Peihong Zhang e Vincent H Crespi. “Smallest nanotube: Breaking the symmetry of sp³ bonds in tubular geometries”. Em: *Physical review letters* 87.12 (2001), p. 125502.
- [17] Xiao-Dong Wen, Roald Hoffmann e NW Ashcroft. “Benzene under high pressure: a story of molecular crystals transforming to saturated networks, with a possible intermediate metallic phase”. Em: *Journal of the American Chemical Society* 133.23 (2011), pp. 9023–9035.
- [18] Martin Olbrich, Peter Mayer e Dirk Trauner. “A step toward polytwistane: synthesis and characterization of C₂-symmetric tritwistane”. Em: *Organic & biomolecular chemistry* 12.1 (2014), pp. 108–112.
- [19] Bo Chen et al. “Linearly polymerized benzene arrays as intermediates, tracing pathways to carbon nanothreads”. Em: *Journal of the American Chemical Society* 137.45 (2015), pp. 14373–14386.
- [20] Can Feng et al. “Morphology-and dehydrogenation-controlled mechanical properties in diamond nanothreads”. Em: *Carbon* 124 (2017), pp. 9–22.
- [21] Haifei Zhan et al. “From brittle to ductile: a structure dependent ductility of diamond nanothread”. Em: *Nanoscale* 8.21 (2016), pp. 11177–11184.
- [22] Haifei Zhan et al. “The morphology and temperature dependent tensile properties of diamond nanothreads”. Em: *Carbon* 107 (2016), pp. 304–309.
- [23] Chia-Chi Chang et al. “A new lower limit for the ultimate breaking strain of carbon nanotubes”. Em: *ACS nano* 4.9 (2010), pp. 5095–5100.

- [24] Mingjie Liu et al. “Carbyne from first principles: chain of C atoms, a nanorod or a nanorope”. Em: *ACS nano* 7.11 (2013), pp. 10075–10082.
- [25] Haifei Zhan et al. “Low interfacial thermal resistance between crossed ultra-thin carbon nanothreads”. Em: *Carbon* (2020).
- [26] Haifei Zhan et al. “Thermal conductivity of a new carbon nanotube analog: The diamond nanothread”. Em: *Carbon* 98 (2016), pp. 232–237.
- [27] Leonid Arturovich Openov e Aleksei Igorevich Podlivaev. “Thermal stability of diamond-like carbon nanothreads”. Em: *JETP letters* 104.3 (2016), pp. 193–196.
- [28] Weikang Wu et al. “Hybrid Structures and Strain-Tunable Electronic Properties of Carbon Nanothreads”. Em: *The Journal of Physical Chemistry C* 122.5 (2018), pp. 3101–3106.
- [29] Zhenzhen Miao et al. “Effects of 3d-transition metal doping on the electronic and magnetic properties of one-dimensional diamond nanothread”. Em: *Chinese Physics B* (2020).
- [30] Zhenzhen Miao et al. “First-principles study on the effects of doping and adsorption on the electronic and magnetic properties of diamond nanothreads”. Em: *Physica E: Low-dimensional Systems and Nanostructures* (2020), p. 113949.
- [31] Taishan Zhu e Elif Ertekin. “Phonons, localization, and thermal conductivity of diamond nanothreads and amorphous graphene”. Em: *Nano letters* 16.8 (2016), pp. 4763–4772.
- [32] Junfeng Gao et al. “Kinetic theory for the formation of diamond nanothreads with desired configurations: a strain–temperature controlled phase diagram”. Em: *Nanoscale* 10.20 (2018), pp. 9664–9672.
- [33] Ke Duan et al. “Diamond nanothread based resonators: ultrahigh sensitivity and low dissipation”. Em: *Nanoscale* 10.17 (2018), pp. 8058–8065.
- [34] Haifei Zhan et al. “Diamond nanothread as a new reinforcement for nanocomposites”. Em: *Advanced Functional Materials* 26.29 (2016), pp. 5279–5283.
- [35] LW Zhang, WM Ji e KM Liew. “Mechanical properties of diamond nanothread reinforced polymer composites”. Em: *Carbon* 132 (2018), pp. 232–240.
- [36] Lu-Wen Zhang et al. “Atomistic Insights into the Tunable Transition from Cavitation to Crazing in Diamond Nanothread-Reinforced Polymer Composites”. Em: *Research* 2020 (2020), p. 7815462.

- [37] Haifei Zhan et al. “High density mechanical energy storage with carbon nanothread bundle”. Em: *Nature communications* 11.1 (2020), pp. 1–11.
- [38] Haifei Zhan et al. “The best features of diamond nanothread for nanofibre applications”. Em: *Nature communications* 8 (2017), p. 14863.
- [39] Pei Wang, Haifei Zhan e Yuantong Gu. “Molecular Dynamics Simulation of Chiral Carbon Nanothread Bundles for Nanofiber Applications”. Em: *ACS Applied Nano Materials* 3.10 (2020), pp. 10218–10225.
- [40] Chengkai Li et al. “Effective Enhancement of a Carbon Nanothread on the Mechanical Properties of the Polyethylene Nanocomposite”. Em: *The Journal of Physical Chemistry C* (2021).
- [41] JFRV Silveira e AR Muniz. “Functionalized diamond nanothreads from benzene derivatives”. Em: *Physical Chemistry Chemical Physics* 19.10 (2017), pp. 7132–7137.
- [42] Pedro G Demingos e Andre R Muniz. “Carbon nanothreads from polycyclic aromatic hydrocarbon molecules”. Em: *Carbon* 140 (2018), pp. 644–652.
- [43] M Leonor Contreras et al. “Exploring the structural and electronic properties of nitrogen-containing exohydrogenated carbon nanotubes: a quantum chemistry study”. Em: *Structural Chemistry* 21.3 (2010), pp. 573–581.
- [44] M Leonor Contreras, Ignacio Villarroel e Roberto Rozas. “How structural parameters affect the reactivity of saturated and non-saturated nitrogen-doped single-walled carbon nanotubes of different chiralities: a density functional theory approach”. Em: *Structural Chemistry* 26.3 (2015), pp. 761–771.
- [45] M Leonor Contreras, Ignacio Villarroel e Roberto Rozas. “Hydrogen physisorption energies for bumpy, saturated, nitrogen-doped single-walled carbon nanotubes”. Em: *Structural Chemistry* 27.5 (2016), pp. 1479–1490.
- [46] Zhuoqun Zheng et al. “Role of Nitrogen on the Mechanical Properties of the Novel Carbon Nitride Nanothreads”. Em: *The Journal of Physical Chemistry C* 123.47 (2019), pp. 28977–28984.
- [47] Min-Min Chen et al. “Theoretical prediction electronic properties of Group-IV diamond nanothreads”. Em: *AIP Advances* 8.7 (2018), p. 075107.
- [48] Julian FRV Silveira e Andre R Muniz. “Diamond nanothread-based 2D and 3D materials: Diamond nanomeshes and nanofoams”. Em: *Carbon* 139 (2018), pp. 789–800.

- [49] Julian FRV Silveira e Andre R Muniz. “Flexible carbon nanothread-based membranes with strain-dependent gas transport properties”. Em: *Journal of Membrane Science* 585 (2019), pp. 184–190.
- [50] Samuele Fanetti et al. “Effect of structural anisotropy in high-pressure reaction of aniline”. Em: *The Journal of Physical Chemistry C* 122.51 (2018), pp. 29158–29164.
- [51] Margaret C Gerthoffer et al. ““Sacrificial” supramolecular assembly and pressure-induced polymerization: toward sequence-defined functionalized nanothreads”. Em: *Chemical Science* 11.42 (2020), pp. 11419–11424.
- [52] Bo Chen et al. “All the Ways To Have Substituted Nanothreads”. Em: *Journal of chemical theory and computation* 14.2 (2018), pp. 1131–1140.
- [53] Wan Si Tang e Timothy A Strobel. “Evidence for Functionalized Carbon Nanothreads from π -Stacked, para-Disubstituted Benzenes”. Em: *The Journal of Physical Chemistry C* 124.45 (2020), pp. 25062–25070.
- [54] Mingrun Du et al. “Vibrational Properties and Polymerization of Corannulene under Pressure, Probed by Raman and Infrared Spectroscopies”. Em: *The Journal of Physical Chemistry C* 123.38 (2019), pp. 23674–23681.
- [55] Haw-Tyng Huang et al. “Nanoarchitecture through Strained Molecules: Cubane-Derived Scaffolds and the Smallest Carbon Nanothreads”. Em: *Journal of the American Chemical Society* (2020).
- [56] Matthew D Ward et al. “Chemistry through cocrystals: pressure-induced polymerization of $C_2H_2 \cdot C_6H_6$ to an extended crystalline hydrocarbon”. Em: *Physical Chemistry Chemical Physics* 20.10 (2018), pp. 7282–7294.
- [57] Yajie Wang et al. “Pressure-Induced Diels-Alder Reactions in C_6H_6 - C_6F_6 Cocrystal towards Graphane Structure”. Em: *Angewandte Chemie International Edition* (2018).
- [58] Matthew D Ward et al. “Controlled Single-Crystalline Polymerization of $C_{10}H_8 \cdot C_{10}F_8$ under Pressure”. Em: *Macromolecules* 52.20 (2019), pp. 7557–7563.
- [59] Srinivasan Marutheeswaran e Eluvathingal D Jemmis. “Adamantane-Derived Carbon Nanothreads: High Structural Stability and Mechanical Strength”. Em: *The Journal of Physical Chemistry C* 122.14 (2018), pp. 7945–7950.

- [60] Alexandra Friedrich et al. “Pressure-Induced Polymerization of Polycyclic Arene–Perfluoroarene Cocrystals: Single Crystal X-ray Diffraction Studies, Reaction Kinetics, and Design of Columnar Hydrofluorocarbons”. Em: *Journal of the American Chemical Society* 142.44 (2020), pp. 18907–18923.
- [61] E. Kaxiras. *Atomic and Electronic Structure of Solids*. Cambridge University Press, 2003.
- [62] R. G. Parr. *Density Functional Theory of Atoms and Molecules*. Oxford University Press, 1989.
- [63] F. Jensen. *Introduction to Computational Chemistry*. Wiley, 2006.
- [64] D. R Hartree. “The wave mechanics of an atom with a non-Coulomb central field. Part I. Theory and methods”. Em: *Proc. Cambridge Philos. Soc.* 24 (1928).
- [65] John C Slater. “Note on Hartree’s method”. Em: *Physical Review* 35.2 (1930), p. 210.
- [66] Vladimir Fock. “Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems”. Em: *Zeitschrift für Physik* 61.1-2 (1930), pp. 126–148.
- [67] Pierre Hohenberg e Walter Kohn. “Inhomogeneous electron gas”. Em: *Physical review* 136.3B (1964), B864.
- [68] Walter Kohn e Lu Jeu Sham. “Self-consistent equations including exchange and correlation effects”. Em: *Physical review* 140.4A (1965), A1133.
- [69] Andreas Görling. “Density-functional theory for excited states”. Em: *Physical Review A* 54.5 (1996), p. 3912.
- [70] Andreas Görling. “Proper treatment of symmetries and excited states in a computationally tractable kohn-sham method”. Em: *Physical review letters* 85.20 (2000), p. 4229.
- [71] Sebastien Hamel et al. “Kohn–Sham orbitals and orbital energies: fictitious constructs but good approximations all the same”. Em: *Journal of electron spectroscopy and related phenomena* 123.2-3 (2002), pp. 345–363.
- [72] Manoj K Harbola. “Relationship between the highest occupied Kohn-Sham orbital eigenvalue and ionization energy”. Em: *Physical Review B* 60.7 (1999), p. 4545.
- [73] Walter Kohn. “Nobel Lecture: Electronic structure of matter—wave functions and density functionals”. Em: *Reviews of Modern Physics* 71.5 (1999), p. 1253.
- [74] John P Perdew, Kieron Burke e Matthias Ernzerhof. “Generalized gradient approximation made simple”. Em: *Physical review letters* 77.18 (1996), p. 3865.

- [75] Stefan Grimme. “Semiempirical GGA-type density functional constructed with a long-range dispersion correction”. Em: *Journal of computational chemistry* 27.15 (2006), pp. 1787–1799.
- [76] Daan Frenkel e Berend Smit. *Understanding Molecular Simulation*. Academic Press, 2002.
- [77] D. C. Rapaport. *The Art of Molecular Dynamics Simulation*. Cambridge University Press, 2004.
- [78] Adri CT Van Duin et al. “ReaxFF: a reactive force field for hydrocarbons”. Em: *The Journal of Physical Chemistry A* 105.41 (2001), pp. 9396–9409.
- [79] Alex D MacKerell Jr et al. “All-atom empirical potential for molecular modeling and dynamics studies of proteins”. Em: *The journal of physical chemistry B* 102.18 (1998), pp. 3586–3616.
- [80] Scott J Weiner et al. “A new force field for molecular mechanical simulation of nucleic acids and proteins”. Em: *Journal of the American Chemical Society* 106.3 (1984), pp. 765–784.
- [81] Huai Sun. “COMPASS: an ab initio force-field optimized for condensed-phase applications overview with details on alkane and benzene compounds”. Em: *The Journal of Physical Chemistry B* 102.38 (1998), pp. 7338–7364.
- [82] William L Jorgensen, David S Maxwell e Julian Tirado-Rives. “Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids”. Em: *Journal of the American Chemical Society* 118.45 (1996), pp. 11225–11236.
- [83] Steven J Stuart, Alan B Tutein e Judith A Harrison. “A reactive potential for hydrocarbons with intermolecular interactions”. Em: *The Journal of chemical physics* 112.14 (2000), pp. 6472–6486.
- [84] J Tersoff. “Empirical interatomic potential for carbon, with applications to amorphous carbon”. Em: *Physical Review Letters* 61.25 (1988), p. 2879.
- [85] Tzu-Ray Shan et al. “Charge-optimized many-body potential for the hafnium/hafnium oxide system”. Em: *Physical Review B* 81.12 (2010), p. 125328.
- [86] Murray S Daw e Michael I Baskes. “Semiempirical, quantum mechanical calculation of hydrogen embrittlement in metals”. Em: *Physical review letters* 50.17 (1983), p. 1285.

- [87] Thomas P Senftle et al. “The ReaxFF reactive force-field: development, applications and future directions”. Em: *npj Computational Materials* 2 (2016), p. 15011.
- [88] Kimberly Chenoweth, Adri CT Van Duin e William A Goddard. “ReaxFF reactive force field for molecular dynamics simulations of hydrocarbon oxidation”. Em: *The Journal of Physical Chemistry A* 112.5 (2008), pp. 1040–1053.
- [89] Tu C Le e David A Winkler. “Discovery and optimization of materials using evolutionary approaches”. Em: *Chemical reviews* 116.10 (2016), pp. 6107–6132.
- [90] Zackary Falls et al. “The XtalOpt Evolutionary Algorithm for Crystal Structure Prediction”. Em: *The Journal of Physical Chemistry C* (2020).
- [91] John Henry Holland. *Adaptation in natural and artificial systems*. MIT Press, 1975.
- [92] I Rechenberg. *Evolutionsstrategie: Optimierung Technischer Systeme nach Prinzipien der Biologischen Evolution*. Frommann-Holzboog, 1973.
- [93] Yanchao Wang et al. “An effective structure prediction method for layered materials based on 2D particle swarm optimization algorithm”. Em: *The Journal of chemical physics* 137.22 (2012), p. 224108.
- [94] D Wolf, OV Buyevskaya e M Baerns. “An evolutionary approach in the combinatorial selection and optimization of catalytic materials”. Em: *Applied Catalysis A: General* 200.1-2 (2000), pp. 63–77.
- [95] David M Deaven e Kai-Ming Ho. “Molecular geometry optimization with a genetic algorithm”. Em: *Physical review letters* 75.2 (1995), p. 288.
- [96] Scott M Woodley e Richard Catlow. “Crystal structure prediction from first principles”. Em: *Nature materials* 7.12 (2008), pp. 937–946.
- [97] Sven Heiles e Roy L Johnston. “Global optimization of clusters using electronic structure methods”. Em: *International Journal of Quantum Chemistry* 113.18 (2013), pp. 2091–2109.
- [98] Tarak K Patra et al. “Defect dynamics in 2-D MoS₂ probed by using machine learning, atomistic simulations, and high-resolution microscopy”. Em: *ACS nano* 12.8 (2018), pp. 8006–8016.

- [99] Lasse B Vilhelmsen e Bjørk Hammer. “A genetic algorithm for first principles global structure optimization of supported nano structures”. Em: *The Journal of chemical physics* 141.4 (2014), p. 044711.
- [100] Artem R Oganov e Colin W Glass. “Crystal structure prediction using ab initio evolutionary techniques: Principles and applications”. Em: *The Journal of chemical physics* 124.24 (2006), p. 244704.
- [101] U Rodemerck et al. “Application of a genetic algorithm and a neural network for the discovery and optimization of new solid catalytic materials”. Em: *Applied Surface Science* 223.1-3 (2004), pp. 168–174.
- [102] William W Tipton e Richard G Hennig. “A grand canonical genetic algorithm for the prediction of multi-component phase diagrams and testing of empirical potentials”. Em: *Journal of Physics: Condensed Matter* 25.49 (2013), p. 495401.
- [103] Silvia Bahmann e Jens Kortus. “EVO—Evolutionary algorithm for crystal structure prediction”. Em: *Computer Physics Communications* 184.6 (2013), pp. 1618–1625.
- [104] Benjamin C Revard et al. “Grand-canonical evolutionary algorithm for the prediction of two-dimensional materials”. Em: *Physical Review B* 93.5 (2016), p. 054117.
- [105] Adam Mrozek, Waław Kuś e Tadeusz Burczyński. “Nano level optimization of graphene allotropes by means of a hybrid parallel evolutionary algorithm”. Em: *Computational Materials Science* 106 (2015), pp. 161–169.
- [106] Nadine L Abraham e Matthew IJ Probert. “A periodic genetic algorithm with real-space representation for crystal structure and polymorph prediction”. Em: *Physical Review B* 73.22 (2006), p. 224104.
- [107] Giancarlo Trimarchi e Alex Zunger. “Global space-group optimization problem: Finding the stablest crystal structure without constraints”. Em: *Physical Review B* 75.10 (2007), p. 104113.
- [108] SQ Wu et al. “An adaptive genetic algorithm for crystal structure prediction”. Em: *Journal of Physics: Condensed Matter* 26.3 (2014), p. 035402.
- [109] Alessandro Fadda e Giuseppe Fadda. “An evolutionary algorithm for the prediction of crystal structures”. Em: *Physical Review B* 82.10 (2010), p. 104105.
- [110] Ask Hjorth Larsen et al. “The atomic simulation environment—a Python library for working with atoms”. Em: *Journal of Physics: Condensed Matter* 29.27 (2017), p. 273002.

- [111] Nathan S Froemming e Graeme Henkelman. “Optimizing core-shell nanoparticle catalysts with a genetic algorithm”. Em: *The Journal of chemical physics* 131.23 (2009), p. 234103.
- [112] Farren Curtis et al. “GAtor: a first-principles genetic algorithm for molecular crystal structure prediction”. Em: *Journal of chemical theory and computation* 14.4 (2018), pp. 2246–2264.
- [113] AN Kolmogorov et al. “New superconducting and semiconducting Fe-B compounds predicted with an ab initio evolutionary search”. Em: *Physical review letters* 105.21 (2010), p. 217003.
- [114] Arunima K Singh et al. “Genetic algorithm prediction of two-dimensional group-IV dioxides for dielectrics”. Em: *Physical Review B* 95.15 (2017), p. 155426.
- [115] Artem R Oganov et al. “Ionic high-pressure form of elemental boron”. Em: *Nature* 457.7231 (2009), pp. 863–867.
- [116] Roy L Johnston. “Evolving better nanoparticles: Genetic algorithms for optimising cluster geometries”. Em: *Dalton Transactions* 22 (2003), pp. 4193–4207.
- [117] SM Woodley e CRA Catlow. “Structure prediction of titania phases: Implementation of Darwinian versus Lamarckian concepts in an Evolutionary Algorithm”. Em: *Computational Materials Science* 45.1 (2009), pp. 84–95.
- [118] Darrell Whitley, Soraya Rana e Robert B Heckendorn. “The island model genetic algorithm: On separability, population size and convergence”. Em: *Journal of computing and information technology* 7.1 (1999), pp. 33–47.
- [119] Paolo Giannozzi et al. “QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials”. Em: *Journal of physics: Condensed matter* 21.39 (2009), p. 395502.
- [120] Python-Software-Foundation. *Python*. 2001. URL: <https://www.python.org/> (acesso em 22/05/2019).
- [121] NumPy-Team. *NumPy*. 2019. URL: <https://www.numpy.org/> (acesso em 22/05/2019).
- [122] Pedro Demingos. *ChemStruct: A Python Package for Structural Analysis of Atomic Systems*. 2019. URL: <https://github.com/pdemingos/chemstruct> (acesso em 22/05/2019).

- [123] Steve Plimpton. “Fast parallel algorithms for short-range molecular dynamics”. Em: *Journal of computational physics* 117.1 (1995), pp. 1–19.
- [124] Kenno Vanommeslaeghe et al. “CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields”. Em: *Journal of computational chemistry* 31.4 (2010), pp. 671–690.
- [125] Jonathan E Mueller, Adri CT Van Duin e William A Goddard III. “Development and validation of ReaxFF reactive force field for hydrocarbon chemistry catalyzed by nickel”. Em: *The Journal of Physical Chemistry C* 114.11 (2010), pp. 4939–4949.
- [126] Claude Charrier et al. “Proton [1, 5] shifts in P-unsubstituted 1H-phospholes. Synthesis and chemistry of 2H-phosphole dimers”. Em: *Journal of the American Chemical Society* 105.23 (1983), pp. 6871–6877.
- [127] Bo Chen, Vincent Crespi e Roald Hoffmann. “Theoretical studies of furan and thiophene nanothreads: structures, cycloaddition barriers and activation volumes”. Em: (2021).