

CONFIABILIDADE DOS METODOS SEMIEMPIROCOS E DFTB PARA O CENÁRIO GLOBAL DE OTIMIZAÇÃO DAS ESTRUTURAS DE NANOCLUSTER

*Reliability of semiempirical and DFTB methods for the global
optimization of the structures of nanoclusters*

Breno R. L. Galvao e Nelson Ribeiro de Carvalho Junior*

*Autor correspondente: nelsonjunior.info@gmail.com

RESUMO

Muitos processos químicos comercialmente e ambientalmente relevantes, com importância crucial para nossa sociedade, podem se beneficiar do desenvolvimento de catalisadores heterogêneos mais ativos ou seletivos. As geometrias de nano-ligas de alumínio e magnésio são exploradas usando um algoritmo genético aplicado a função potencial Gupta e sintonizada para procurar os mínimos de energia mais baixos para cada tamanho de cluster. Na busca por mínimos disponíveis nas hipersuperfícies de energia potencial associadas aos clusters são importantes para entender catalisadores de átomos únicos em superfícies de metais nobres, nos quais a diferença em seletividade e atividade entre o metal hospedeiro e os dopantes é utilizada, verificou-se que as energias encontradas em muitos casos ficaram estagnadas não havendo melhoras; Quando os resultados obtidos da geração anterior não evoluem foi acrescentado um novo fluxo onde o número da geração é avaliado podendo alterar o valor do gradiente de busca local o que torna o Algoritmo Genético mais consistente em busca do mínimo global.

Palavras-chave: Potencial Gupta, Energia, Algoritmo Genético, Alumínio, Magnésio.

ABSTRACT

Many commercially and environmentally relevant chemical processes of crucial importance to our society could benefit from the development of more active or selective heterogeneous catalysts. The geometries of aluminum and magnesium nano-alloys are explored using a genetic algorithm applied to the Gupta potential function and tuned to look for the lowest energy minima for each cluster size. important to understand single atom catalysts on noble metal surfaces, in which the difference in selectivity and activity between the host metal and dopants is used, it was found that the energies found in many cases remained stagnant with no improvements; When the results obtained from the previous generation do not evolve, a new flow was added where the generation number is evaluated and can change the value of the local search gradient, which makes the Genetic Algorithm more consistent in search of the global minimum.

Keywords: Potential Gupta, Energy, Genetic Algorithm, Aluminum, Magnesium

1. Introdução

O potencial oferecido pela atual tecnologia permitiu o aprimoramento de uma grande variedade de técnicas numéricas, que exigem um grande esforço computacional. Esses aprimoramentos possibilitaram o desenvolvimento de muitas áreas de aplicação, entre elas a Química Teórica e Computacional que associada ao advento da nanotecnologia torna-se um dos domínios interdisciplinares mais promissores do século XXI.

A nanotecnologia resulta de estudos avançados nas áreas de Física, Química, Biologia, Computação e Engenharia de Materiais. A troca de experiências e conhecimentos no campo da chamada nanociência permite a construção de instrumentos e técnicas capazes de manipular átomos e moléculas de diversos materiais, acrescentando-lhes novas propriedades e abrindo caminho para as mais diversas aplicações. E entre elas destaca-se o uso dos Algoritmos Genéticos para encontrar o mínimo global e os mínimos locais de energia mais baixa.

O uso de Algoritmos Genéticos para a otimização da geometria de agregados teve início na década de 90, para a otimização de pequenos agregados de silício e agregados moleculares, respectivamente [Chen, 2013]. A otimização da geometria de agregados é um processo que busca encontrar a configuração mais estável para um composto iônico. Em ambos os casos, a geometria dos agregados era codificada por bits. Em seguida, Deaven e Ho desenvolveram um algoritmo híbrido onde se aplica uma minimização local por gradiente sempre que uma nova estrutura é gerada pelo Algoritmo Genético [Aguado, 2010].

O alumínio por suas excelentes características físico químico é empregado na indústria civil, automotiva, naval e aeroespacial. As características mecânicas dos sistemas solidificados dependem do arranjo micro estrutural, mais especificamente do tamanho de grão, espaçamentos celulares ou dendríticos, das heterogeneidades de composição química, do tamanho, forma, distribuição das inclusões e das porosidades. [Parr, 1989]

2. Metodologia

2.1. Potencial Empírico

A função analítica empírica Gupta foi utilizada visando uma avaliação rápida da energia potencial do sistema como método de minimização durante as

buscas do Algoritmo Genético.

Este potencial é amplamente utilizado no tratamento de sistemas metálicos, as ligações metálicas ocorrem entre átomos de metais, caracterizadas por uma "nuvem" de elétrons livres que se movem entre os átomos. Esse fenômeno ocorre porque os metais possuem elétrons de valência que não estão fortemente ligados aos seus núcleos, permitindo que se desloquem com facilidade. onde a energia de ligação total de um cluster composto de N átomos, $E_b(N)$, é escrito como uma soma em termos de energia atômica decompostas nos termos como segue abaixo:

$$E_i = E_i^{\text{band}} + E_i^{\text{rep}} \quad (1)$$

$$E_i^{\text{band}} = - \left\{ \sum_{j \neq i}^N \zeta^2 \exp \left[-2q \left(\frac{r_{ij}}{r_0} - 1 \right) \right] \right\}^{1/2} \quad (2)$$

$$E_i^{\text{rep}} = \sum_{j \neq i}^N A \exp \left[-p \left(\frac{r_{ij}}{r_0} - 1 \right) \right] \quad (3)$$

O termo E^{band} é uma expressão para muitos corpos e compõe a parte atrativa da energia, levando em consideração o caráter de banda da ligação metálica [Chen, 2013].

Por outro lado, deve haver uma contribuição repulsiva para assegurar a estabilidade da estrutura, que é introduzida como a interação entre pares atômicos que compõe o segundo termo da energia de coesão [Chen, 2014]. Eles são escritos como uma função da distância entre os átomos (r_{ij}), e os parâmetros livres é determinada pelo ajuste do potencial para resultados experimentais de energia de coesão, parâmetros de rede e constantes elásticas [Schuchardt, 1998].

Fisicamente, E e A determinam as forças dos termos atraentes e repulsivos e têm unidades de energia, enquanto r_0 representa a distância do vizinho mais próximo.

Para as interações Mg-Mg, esses parâmetros foram retirados de um trabalho recente de [Chen, 2014] e foram relatados para melhorar a predição dos resultados experimentais.

Os parâmetros para alumínio foram obtidos a partir da ref. [Weigend, 2005] enquanto aqueles para Al-Mg foram obtidas tomando a média aritmética dos homo-nucleares.

Tal potencial empírico é mais eficiente na descrição de grandes aglomerados, uma vez que eles são ajustados a parâmetros. Para pequenas estruturas,

diferentes efeitos quânticos podem dominar, e a função analítica simples pode produzir estruturas incorretas como mínimos globais. No entanto, foi demonstrado que mesmo em pequenas escalas, tais como oito átomos de clusters, o potencial empírico é geralmente um bom ponto de partida e tende a melhorar para maiores clusters [Stephens, 1994]

Na busca por mínimos disponíveis função de energia potencial associadas aos clusters estudados, o método de minimização escolhido, em foi o algoritmo genético (GA) [Goldberg,1989].

Conforme os clusters crescem, o número de mínimos locais presente na energia potencial aumenta drasticamente, mascarando o mínimo global e tornando a tarefa de encontrá-lo muito mais laboriosa.

Nesta abordagem, para um cluster de um dado tamanho e de uma dada composição, uma população inicial gerada aleatoriamente (através da geração aleatória das coordenadas de cada átomo dentro de um dado intervalo). Em seguida, suas energias são calculadas pelo potencial Gupta e o passo extra de permitir que as estruturas relaxem para o mínimo local mais próximo é dado utilizando o método DFT de minimização baseado no cálculo do vetor gradiente, este cálculo é realizado considerando um erro muito pequeno, que quase se assemelha a uma busca local a cada ponto da população de cada geração.

Os indivíduos são então ranqueados segundo suas estabilidades energéticas de acordo com uma função de adequação chamada *fitness function*, que no presente trabalho foi escolhida pela expressão apresentada abaixo.

$$F_i = \frac{1}{2} \left[1 - \tanh \left(2 \frac{E_i - E_{\min}}{E_{\max} - E_{\min}} - 1 \right) \right] \quad (4)$$

Onde:

$E_{\max}$ -- O máximo da energia potencial da população.

$E_{\min}$ -- O mínimo energia potencial da população

Os operadores evolucionários do procedimento padrão do GA (cruzamento, mutação e seleção natural) são então aplicados à população, onde os indivíduos progenitores são selecionados para cruzamento através de uma variante do método da roleta para fornecer nova descendência.

O método da roleta seleciona os progenitores em uma população fazendo com que a probabilidade de um cluster ser escolhido para cruzamento seja proporcional ao seu valor de adequação dado pela

função fitness.

Os resultados obtidos a cada geração são analisados conforme ilustrado no fluxograma da figura 1, onde o critério de análise foi verificar a cada duas gerações a energia gerada pelo potencial empírico Gupta e confrontada com as gerações anteriores, caso não houver evolução da energia o valor do gradiente (valor do método DFT que está variando de 0,01 a 0,000001) é alterado possibilitando tornar o GA mais consistente em busca do mínimo global, onde a iteração do valor gradiente é definido:

$\max. \{ |proj g_i| \mid i = 1, \dots, n \} \leq gtol_{\max} \{ |proj g_i| \mid i = 1, \dots,$

$n \} \leq gtol$

Onde: pg_i é o i -ésimo componente do gradiente projetado.

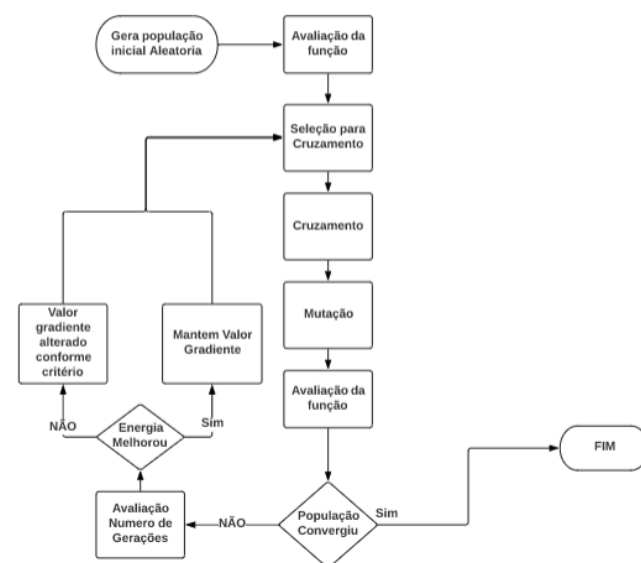


Figura 1. Fluxograma descrevendo a abordagem evolutiva com dois novos processos para otimização dos ciclos do GA.

3. Resultados

Realizamos os procedimentos do GA descrito acima para aglomerados de Al_xMg_y com um orçamento fixo (critério de parada de 4.000.000 de chamadas ao Gupta para cada semente), experimentos das diversas composições $x+y = 13, 14, 15$ e 20 foram alternando o valor do gradiente = $Gtol$ a medida que não houver melhora na energia conforme os novos fluxos

permitiram o controle da variação do gradiente entre 0,01 até 0,000001.

A proposta dos experimentos é analisar este novo modelo possibilitando explorar todo potencial do Algoritmo Genético, analisando seu comportamento para pequenas e médias composições a cada ciclo, os experimentos foram divididos em duas etapas conforme resultados apresentados a seguir.

3.1. Composições diversas

Os experimentos desta seção foram denominados de 1ª etapa onde foi essencial para validação deste novo metodo, foi analisado o comportamento do Algoritmo Genético para diversas composições e valores de Gradiente alternados, foi levado em consideração para análise, a média da energia obtida em XX execuções independentes e o desvio padrão de cada metodo. Nos experimentos o Algoritmo Genético foi classificados de:

- Sequencial – Tradicional em vários experimentos na literatura onde o Potencial Empirico Gupta é chamado a todo momento com o valor do gradiente fixo correspondente ao metodo de otimização DFT de 0,000001, valor definido pela *Scientific Computing with Python*.
- Alternado – Criado para alterar o valor do Gradiente conforme descrito no fluxograma figura 1 onde é considerado o número de gerações corrente e analisado se houve melhora na evolução das energias, comparando com a energia gerada a ciclo e conforme criterio utilizado foi alterado o valor do Gradiente de 0,01 até 0,000001.

Todos os experimentos desta primeira etapa foram realizados com 10 sementes com composições variando de $x+y = 15$ a $x+y = 20$

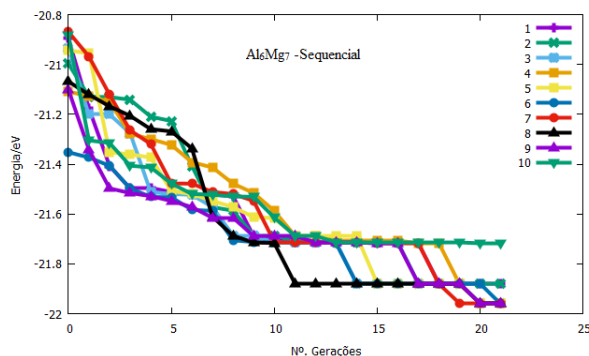


Figura 2. Composição Al_6Mg_7 com chamadas sequenciais ao método de otimização DFT para 10 execuções independentes utilizando do algoritmo.

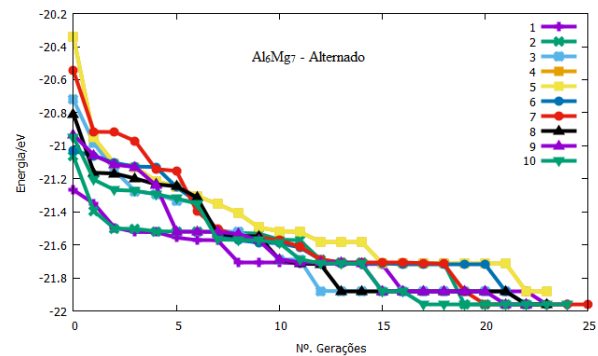


Figura 3. Composição Al_6Mg_7 com chamadas Alternadas ao método de otimização DFT para 10 execuções independentes utilizando do algoritmo.

Na tabela 1 observa-se que o método alternado mostrou-se eficaz para a composição Al_6Mg_7 , considerada uma composição pequena com 13 átomos, sendo melhor no desvio padrão, media da energia e obtendo uma melhor exploração do Algoritmo Genético com 4 gerações mais do que o método sequencial. Abaixo outro experimento com uma composição diferente e com um maior número de átomos $x+y = 20$.

Tabela 1. Resultados da composição Al_6Mg_7 com 10 sementes

Parâmetros	Sequencial	Alternado
Nº. Gerações	21	25
Media da Energia/eV	-21,9198158950	-21,9441952625
Desvio Padrão	0,07884268	0,03376330

No próximo experimento foi realizado com uma composição maior de átomos do que os experimentos anteriores sendo $x+y = 20$.

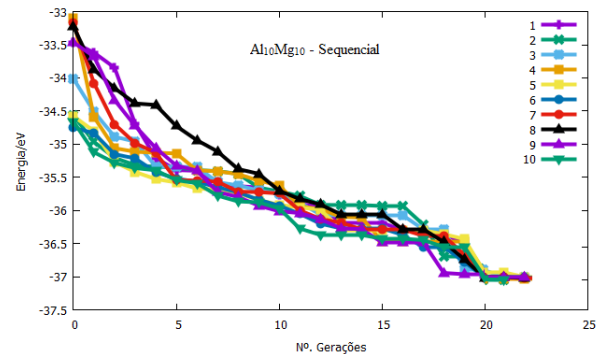


Figura 4. Composição $Al_{10}Mg_{10}$ com chamadas sequenciais ao método de otimização DFT para 10 execuções independentes utilizando do algoritmo.

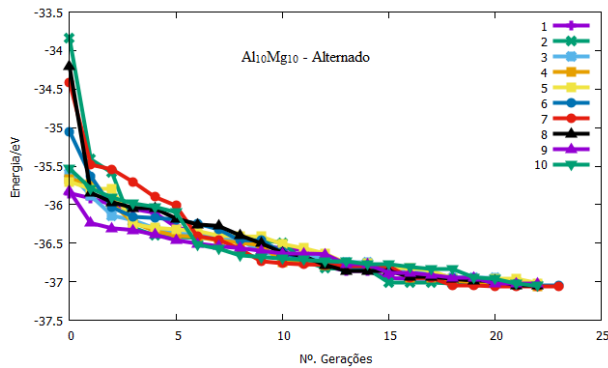


Figura 5. Composição Al10Mg10 com chamadas alternadas ao método de otimização DFT para 10 execuções independentes utilizando do algoritmo.

Na tabela 2 ilustra os resultados obtidos dos experimentos com uma composição maior do que os experimentos anteriores, mas mesmo com esta composição maior o comportamento do GA continuou sendo melhor na média da energia, no desvio padrão e obtendo uma maior exploração do GA alcançando 1 geração a mais do que o método sequencial.

Tabela 2. Resultados da composição Al10Mg10 com 10 sementes

Parâmetros	Sequencial	Alternado
Nº. Gerações	22	23
Media da Energia/eV	-37,03828549	-37,03892158
Desvio Padrão	0,0257267	0,015848

Para comparar a eficácia deste novo método foi realizada uma segunda etapa com o mesmo número de átomos Al10Mg10 mais com o triplo das sementes buscando analisar o comportamento do GA para um orçamento ainda maior do critério de parada.

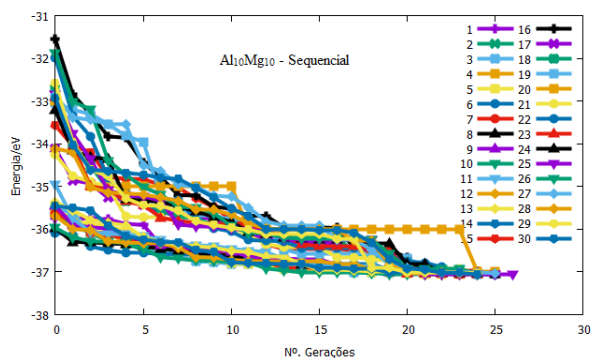


Figura 6. Composição Al10Mg10 com chamadas sequenciais ao método de otimização DFT para 30 execuções independentes utilizando do algoritmo.

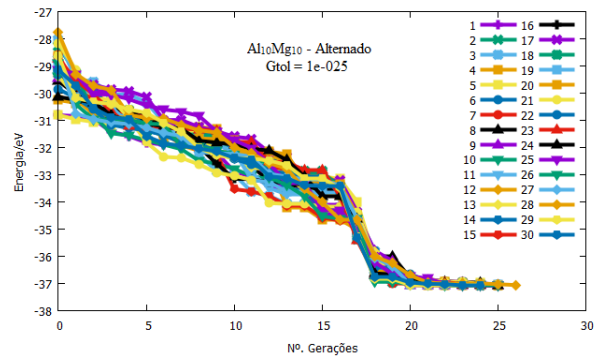


Figura 7. Al10Mg10 com chamadas alternadas ao método de otimização DFT para 30 execuções independentes utilizando do algoritmo.

Na tabela 3 ilustra os resultados obtidos através dos experimentos realizados com 30 sementes, cujo os mesmos foram bastante promissores onde o GA ficou com a melhor média de energia e foi acrescentado um quarto parâmetro para validar ainda mais este novo método, que a quantidade de vezes que o GA encontrou mínimo global.

Tabela 5. Resultados da composição Al10Mg10 com 30 sementes

Parâmetros	Sequencial	Alternado
Nº. Gerações	25	25
Media da Energia/eV	-37,03828549	-37,03892158
Desvio Padrão	0,02137525	0,02663079
Nº Mínimo global	9	10

A figura 8 ilustra o comportamento do desvio padrão dos experimentos realizados, onde notas claramente que o desvio padrão dos Algoritmos Genéticos foi melhor em comparação ao Algoritmo Genético Sequencial.

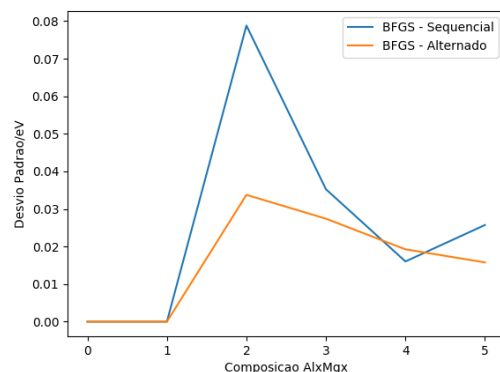


Figura 8. Comportamento do desvio padrão dos experimentos realizados.

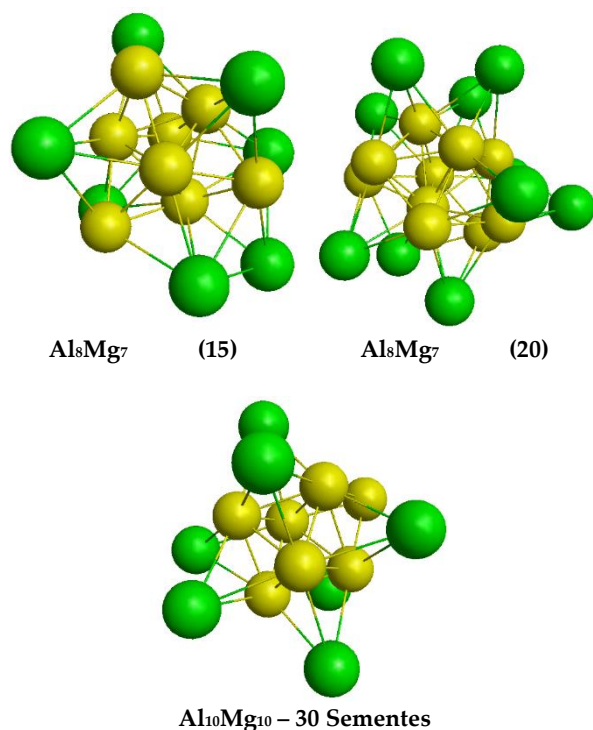


Figura 19. Moléculas Geradas pelos experimentos

4. Conclusões

O presente artigo teve por objetivo apresentar um novo método para explorar DFT onde foi acrescentado um novo fluxo de tomada de decisão em busca de otimizar e melhorar computacionalmente o desempenho do Algoritmo Genético, controlando o valor do gradiente conforme sua evolução. Este novo método mostrou-se muito promissor de acordo com resultados apresentados na seção anterior, uma vez que o Algoritmo Genético Sequencial que se torna bastante custoso pois invoca o método de otimização DFT sempre com erro do vetor Gradiente muito pequeno, diferentemente do Algoritmo Genético que começa com um valor bem baixo e vai aumentando gradativamente conforme critérios estabelecidos. Nos experimentos realizados tanto com 10 sementes quanto com 30 sementes o Algoritmo Genético obteve melhores energias na média geral e também o desvio padrão ficou bem melhor, alcançando melhores mínimos globais.

Realizar mais experimentos com composições bem maiores para verificarmos a eficácia e o comportamento deste novo método de otimização.

Referências

Aguado, A. & López, J. M. J. Chem. (2010), Phys, 133, 094302.

Becke, A. D. J. Chem. (1993), Phys. 98, 5648.

Chen, Z., Jiang, X., Li, J. & Li, S. J. Chem. (2013), Phys. 138, 214303.

Chen, C. et al. Science (2014), Wilcoxon, J. P. & Provencio, P. P. J. Am. Chem. Soc. 126, 6402.

D.E.Goldberg (1989). Genetic Algorithms in Search Optimization and Machine Learning Addison Wesley Longman, Inc.

Parr, R. G. & Yang, (1989), W. Density-Functional Theory of Atoms and Molecules. (Oxford University Press,).

Schuchardt, (1998), K. L. et al. J. Chem. Inf. Model.

Stephens, (1994), P. J., Devlin, F. J., Chabalowski, C. F. & Frisch, M. J. J. Phys. Chem. 98, 11623.

Tchaplyguine, (2009), M. et al. Phys. Rev. B: Condens. Matter Mater. Phys. 80, 033405

Weigend, F. & Ahlrichs, R. (2005). Phys. Chem. Chem. Phys. 7, 3297.

Watts, J. D., Gauss, J. & Bartlett, (1993). R. J. J. Chem. Phys. 98, 8718-8733 (1993).