

Dinâmica molecular da adenosina deaminase de *Edwardsiella tarda* em água - Tutorial experimental

Paulo Henrique Matayoshi Calixto¹, Arnaldo José Ballarini², Claudio Alberto Gellis de Mattos Dias³, Júlio César Sá-Oliveira⁴

1. Biomédico, Mestre e Doutorando em Medicina Tropical e Infectologia, Sub-área: Parasitologia e Imunologia Aplicadas, Universidade Federal do Triângulo Mineiro, Brasil. E-mail: ph_calixto@yahoo.com

2. Médico Veterinário, Universidade Federal Rural do Rio de Janeiro. Mestre em Desenvolvimento Regional, Universidade Federal do Amapá. Professor, Universidade Federal do Amapá, Brasil. E-mail: ballarini@unifap.br

3. Biólogo, Universidade do Sagrado Coração. Doutor em Teoria da Pesquisa do Comportamento, Universidade Federal do Pará. Professor Efetivo, Instituto Federal do Amapá, Brasil. E-mail: claudioidias@unifap.br

4. Biólogo, Universidade Federal do Rio Grande do Norte. Doutor em Ecologia Aquática e Pesca, Universidade Federal do Pará. Professor Efetivo, Universidade Federal do Amapá. Coordenador do Núcleo de Estudos em Pesca e Aquicultura-UNIFAP, Brasil. E-mail: juliosa@unifap.br

RESUMO: A bactéria Gram-negativa *Edwardsiella tarda* é responsável pela edwardsielose, uma doença hemorrágica de peixes, que conduz a grandes perdas econômicas em todo o mundo. O conhecimento das moléculas envolvidas em processos bioquímicos de *E. tarda* é importante para o desenvolvimento de novos alvos terapêuticos e/ou profiláticos. Uma dessas moléculas é a adenosina deaminase, envolvida no processo de maturação do tRNA, através da conversão de adenosina em inosina, ainda não caracterizada em *E. tarda*. A caracterização de uma proteína é mais bem conduzida quando se conhece as suas particularidades estruturais. Uma das principais ferramentas de caracterização estrutural é a dinâmica molecular. Nesse sentido, o objetivo deste trabalho foi de proporcionar um tutorial facilitado de dinâmica molecular utilizando o GROMACS, bem como, estabelecer a dinâmica molecular *in silico* da adenosina deaminase de *E. tarda*. Os dados gerados apontaram que a estrutura da adenosina deaminase é estável, tanto do ponto de vista estrutural, quanto de enovelamento. O uso da dinâmica molecular pode auxiliar na escolha de novos alvos terapêuticos, sobretudo, no desenho racional de fármacos.

Palavras-chave: *Edwardsiella tarda*, adenosina deaminase, dinâmica molecular, GROMACS, tutorial.

Molecular dynamics of *Edwardsiella tarda* adenosine deaminase protein in water - Experimental tutorial

ABSTRACT: The Gram-negative bacteria *Edwardsiella tarda* is responsible for edwardsiellosis, a hemorrhagic disease of fish, which leads to large economic losses worldwide. The knowledge of the molecules involved in biochemical processes in *E. tarda* is important for the development of new therapeutic and/or prophylactic targets. One of these molecules is adenosine deaminase, involved in the maturation process of the tRNA by converting adenosine to inosine, not yet investigated in *E. tarda*. The characterization of a protein is best conducted when if know the structural particularities. One of the main tools of structural characterization is molecular dynamics. In this sense, the aim of this work was provide to the reader an eased tutorial of molecular dynamics using the GROMACS, as well as perform molecular dynamics of *E. tarda* adenosine deaminase. The data generated by molecular dynamics showed that the structure of adenosine deaminase, determined *in silico*, is extremely stable. The use of molecular dynamics can assist in the choice of new therapeutic targets, especially in rational drug design.

Keywords: *Edwardsiella tarda*, adenosine deaminase, molecular dynamics, GROMACS, tutorial.

1. Introdução

A *Edwardsiella tarda*, um bacilo Gram-negativo, é o agente etiológico da edwardsielose. A edwardsielose é uma infecção grave, notadamente conhecida pelos fenômenos de sepsse hemorrágica em peixes (HU et al., 2014). A infecção por *E. tarda* não se limita ao âmbito veterinário, podendo também infectar o homem (JANDA; ABBOTT, 1993). A infecção humana não apresenta fenômenos hemorrágicos e se caracteriza por gastroenterites autolimitadas e em diferentes graus de gravidade (JARURATANASIRIKUL; KALNAUWAKUL, 1991). Mesmo com a existência de vacinas experimentais e de tratamentos baseados em antibióticos, a edwardsielose promove grandes perdas econômicas, seja pelo acometimento de peixes em ambientes naturais ou em criadouros artificiais (WANG et al., 2012).

Embora existam várias vacinas experimentais, nenhuma delas é comercializada, indicando que seu papel protetivo ainda não atingiu o patamar desejado (NELSON et al., 2009). O uso de antibióticos na água acarreta uma série de transtornos, especialmente no que se refere à resistência aos fármacos (ZHENG et al., 2012). Dessa maneira, é mister o desenvolvimento de novas vacinas e/ou fármacos. Nesse sentido, o conhecimento da maquinaria bioquímica da bactéria é de suma importância na escolha de macromoléculas, tais como proteínas, como alvos para o tratamento e/ou profilaxia da

infecção. Uma dessas moléculas é a adenosina deaminase, responsável pela maturação da molécula de RNA transportador (tRNA) através da deaminação da adenosina em inosina, essencial à vida (LUO; SCHRAMM, 2008). Mesmo sendo uma molécula potencialmente importante como alvo para medidas terapêuticas e/ou profiláticas, nenhum estudo sobre a adenosina deaminase foi empreendido em *E. tarda*.

A caracterização de uma proteína é melhor conduzida de posse da estrutura tridimensional (NADEAU; CARLSON, 2012). A partir da estrutura é possível inferir e mensurar o mecanismo de funcionamento proteico e propor desde o aumento da atividade, através de mutações sítio-dirigidas, até a cessação da atividade proteica, por intermédio do desenho racional de fármacos (ELLEBY et al., 2000; KUMAR et al., 2014). Indubitavelmente, uma das melhores ferramentas de caracterização estrutural é a dinâmica molecular; estudo amplamente difundido no campo da Biologia Computacional e em outras áreas de conhecimento *in silico* (KAUSHIK et al., 2014).

As proteínas adotam estruturas tridimensionais baseadas, entre outros fatores, na sequência de aminoácidos (GUZMAN; GRUEBELE, 2014). No entanto, as proteínas não são entidades estritamente estáticas e apresentam diversos graus de movimentos e, conseqüentemente, diferentes conformações (HUB; GROOT, 2009; CHENG; DAGGETT, 2014). Nesse sentido, a dinâmica molecular vem elucidar as mudanças de conformação

proteica, trazendo a tona informações valiosas no que diz respeito às características funcionais da proteína estudada. Por definição, a dinâmica molecular é uma simulação computacional dos movimentos físicos das moléculas em nível atômico em meios sólidos, líquidos ou gasosos (JANSSON et al., 2009). Existem vários programas disponíveis para dinâmica molecular de biomoléculas, sendo os mais empregados na pesquisa o AMBER (WANG et al., 2004), NAMD (PHILLIPS et al., 2005), GROMOS (DAURA et al., 1996), CHARMM (ROTERMAN et al., 1989) e GROMACS (LANGE et al., 2006). Indiscutivelmente, este último é o mais difundido e utilizado em pesquisas que envolvam dinâmica molecular de proteínas. Esse fato se deve aos amplos recursos de parametrização e pela gratuidade do uso para fins acadêmicos (VAN DER SPOEL et al. 2005). No entanto, muitos usuários, especialmente aqueles da área de Ciências Biológicas e da Saúde, apresentam dificuldades em se introduzirem no uso do GROMACS. Isso se deve ao fato do programa não apresentar interface gráfica, sendo cada comando inserido através de linhas de comando num terminal, exigindo prévio conhecimento de informática, sobretudo no sistema operacional Linux. Diante desse cenário, os objetivos deste trabalho são de proporcionar um tutorial detalhado e facilitado do uso do GROMACS, bem como, de realizar o estudo da dinâmica da adenosina deaminase de *E. tarda*.

2. Materiais e métodos

Determinação da estrutura da adenosina deaminase

A estrutura da adenosina deaminase de *E. tarda* (EtAD) (GenBank: WP_015683243.1) (TEKEDAR et al., 2013) foi determinada por modelagem comparativa, através do método de restrições espaciais, empregando o software Modeller 9.13v (SALI et al., 1995). Inicialmente, foram gerados 100 modelos, sendo o que apresentou melhor perfil estereoquímico, eleito para este trabalho.

O refinamento da estrutura foi realizado através da minimização de energia, empregando o campo de força OPLS (Optimized Potentials for Liquid Simulations) (MACCALLUM; TIELEMAN, 2003) A predição dos rotâmeros foi realizada com o auxílio da interface gráfica UCSF Chimera (PETTERSEN et al., 2004), com o uso da biblioteca de rotâmeros de Dunbrack (DUNBRACK, 2002). Após o refinamento, a estrutura foi validada através dos seguintes programas: Procheck (LASKOWSKI et al., 1996), ProSa (WIEDERSTEIN; SIPPL, 2007) e Verify3D (EISENBERG et al., 1997).

Dinâmica molecular

No estudo de dinâmica molecular foi utilizado o software GROningen MACHine for Chemical Simulations – GROMACS (release 4.5.6) (LANGE et al., 2006). A simulação se dividiu em cinco etapas: i) geração da topologia; ii) criação do box, solvatação e adição de íons; iii) minimização de energia; iv) equilíbrio; v) dinâmica molecular de 1 ns. Os parâmetros empregados na simulação são descritos no tutorial.

Aporte de hardware e software

Neste trabalho foram utilizados os seguintes softwares: UCSF Chimera (PETTERSEN et al., 2004) disponível em: <<https://www.cgl.ucsf.edu/chimera/>>, Xmgrace, disponível em: <<http://plasma-gate.weizmann.ac.il/Grace/>> e GROMACS 4.5.6v, disponível em: <<http://www.gromacs.org/>>. Todos os softwares foram executados no sistema operacional Linux, devido sua robustez, estabilidade e eficiência. Caso o usuário tenha dificuldade em instalar os programas, o mesmo poderá instalá-los com o auxílio do Central de Programas do Ubuntu.

Este tutorial foi conduzido em duas máquinas não paralelizadas. A configuração da primeira máquina é a seguinte: processador Intel Core i7 4770K, 32Gb de memória RAM (1600 MHz), HD de 2Tb e GPU VGA GigaByte GeForce GTX660 Ti 2048MB, com sistema operacional Red Hat Linux. Já a segunda máquina, utilizada como suporte, possui o processador FX-8350, 8Gb de memória RAM (1600 MHz), disco rígido de 1Tb de armazenamento e GPU VGA GigaByte GeForce GTX660 Ti 2048MB, com sistema operacional Linux, distribuição Ubuntu 12.04.

3. Resultados e discussão

Obtenção da estrutura e geração da topologia

Como alvo para a DM foi utilizada a estrutura da adenosina deaminase (GenBank: WP_015683243.1) de *E. tarda* (TEKEDAR et al., 2013), resolvida por modelagem comparativa (Fig. 1A). Na impossibilidade do usuário determinar a estrutura *in silico*, deverá ser usada a estrutura da enzima adenosina deaminase (PDB ID: 3OCQ) de *Salmonella enterica* (SeAD) (Fig. 1B), disponível no Protein Data Bank (PDB), website <www.rcsb.org>. Caso seja utilizada a estrutura 3OCQ, o usuário deverá carregar o arquivo 3OCQ.pdb, preferencialmente, no editor de texto Gedit e realizar a retirada das moléculas de água (HOH), bem como do átomo de zinco (Zn) do arquivo .pdb. Esse procedimento não pode ser aplicado a todas as situações, pois algumas moléculas de água, bem como de outros heteroátomos (por exemplo, Zn²⁺, Ca²⁺, Cu²⁺, Mg²⁺, etc.) são essenciais para o processo catalítico de determinadas enzimas. Contudo, para este propósito, nenhuma molécula de água presente no arquivo .pdb será necessária.

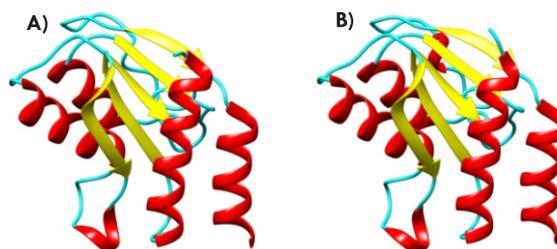


Figura 1. Estruturas tridimensionais da Adenosina deaminase. A) EtAD resolvida por modelagem comparativa. B) SeAD (PDB: 3OCQ) determinada por difração de Raios-X. / **Figure 1.** Adenosine deaminase of three-dimensional structures. A) EtAD resolved by comparative modeling. B) SEAD (PDB: 3OCQ) determined by X-ray diffraction.

É imperioso destacar que, ao recuperar estruturas do Protein Data Bank, o usuário deverá checar o arquivo .pdb em busca de comentários com o termo “MISSING”, pois esta entrada se refere aos átomos ou resíduos ausentes no cristal. A ausência de resíduos nas regiões amino- e carboxi-terminal, a priori, não gera inconvenientes significativos no processo de dinâmica molecular. Por outro lado, quando os resíduos faltantes se localizam na região interna da proteína, a dinâmica pode falhar. Esse obstáculo é facilmente superado através da modelagem do trecho ausente, seja por modelagem comparativa, *threading* ou *ab initio*. Além de checar o arquivo .pdb, o usuário também deverá realizar a inspeção visual da estrutura antes de iniciar a simulação. A verificação visual pode ser realizada com softwares de visualização, tais como UCSF Chimera (PETTERSEN et al., 2004) e VMD (HUMPHREY et al., 1996).

É necessário que o usuário crie a pasta “tutorial”, onde deverá ser armazenada a estrutura da adenosina deaminase. O arquivo .pdb armazenado na pasta “tutorial” será o arquivo de entrada para a ferramenta pdb2gmx do GROMACS. O usuário deverá, através de um terminal, acessar a pasta “tutorial” e inserir a seguinte linha de comando “**pdb2gmx -f 1AKI.pdb -o 1AKI_processed.gro -water spce**”. Após a execução desse

comando, o GROMACS solicitará a escolha do campo de força. Ao realizar outras análises, o usuário deverá ler sobre as particularidades e aplicabilidade de cada campo de força. Para este tutorial, foi escolhido o campo de força OPLS-AA\L (XU et al., 2007), determinado pela opção 14. O `pdb2gmx`, ao fim do processamento, gerará três arquivos, como descrito a seguir: i) `EtAD_processed.gro`, arquivo de estrutura que contém todos os átomos definidos dentro do campo de força escolhido; ii) `topol.top`, apresenta uma descrição completa entre a interação de todas as partículas do sistema e traz informações importantes, tais como: (nr) número do átomo, (type) tipo do átomo, (resnr) número do resíduo, (residue) nome do resíduo, (atom) nome do átomo, (cgnr) número do grupo da carga, (charge) carga, (mass) massa, (typeB, chargeB e massB) usados em cálculos de *free-energy*; iii) `posre.itp`, contém informações sobre a restrição de posição dos átomos pesados.

Definição do box, solvatação e adição de íons

Após a geração do arquivo de topologia (`topol.top`) da estrutura é necessário a realização de três passos fundamentais para o processamento da simulação, são eles: i) delimitação do tamanho e forma do recipiente (*box*) que comportará a proteína; ii) escolha e adição do solvente; iii) verificação da necessidade da adição de íons para alcançar a neutralidade do sistema. A criação do *box* é realizada através do `editconf`, com a seguinte linha de comando, acessando o terminal na pasta tutorial: **`editconf -f EtAD_processed.gro -o EtAD_newbox.gro -c -d 1.0 -bt cubic`**. O comando `-c` conduz a estrutura para o centro do *box*, enquanto o comando `-d 1.0` posiciona a estrutura a 1 nm a partir das bordas. A distância de 1 nm é suficiente para que a estrutura não interaja com sua imagem periódica. Já forma do *box* é definida como cúbica, através do comando `-bt cubic`. Várias formas podem ser definidas, tais como o dodecaedro. Essa forma apresenta vantagens, em termos de dimensões, pois seu volume corresponde a aproximadamente 70% do volume do cubo.

Uma vez definida as dimensões do *box*, é necessário preenche-lo com o solvente (Fig. 2). A adição do solvente é realizada com a seguinte linha de comando: **`genbox -cp EtAD_newbox.gro -cs spc216.gro -o EtAD_solv.gro -p topol.top`**. O comando `-cp` faz menção à configuração da proteína e corresponde ao *output* do `editconf` usado anteriormente. O modelo da água utilizado neste estudo é o SPC (*Simple Point Charge water*), sendo este indicado pelo arquivo `spc216.gro` do GROMACS, através do comando `-cs`. Por fim, a linha de comando do `genbox` atualizará o arquivo de topologia da estrutura.

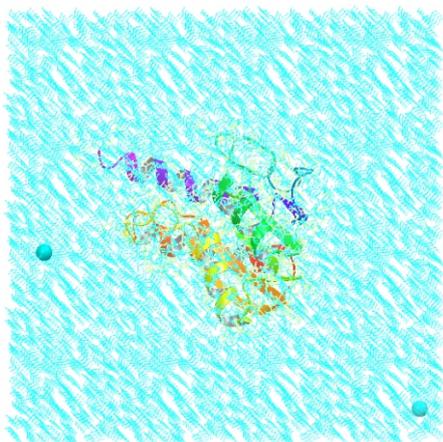


Figura 2. Estrutura da EtAD contida no *box* preenchido com água. As esferas azuis correspondem aos dois íons cloro adicionados para atingir a neutralidade do sistema. / **Figure 2.** Structure EtAD contained in the *box* filled with water. The blue beads correspond to two chlorine ions added to achieve neutrality system.

O sucesso da dinâmica molecular depende, entre outros fatores, da neutralidade do sistema. Esta é alcançada através da adição de íons ao sistema. Por exemplo, caso a carga do sistema seja positiva, adiciona-se íons cloro. No entanto, nos casos onde a carga é negativa, realiza-se a adição de íons sódio. A adição de íons pelo GROMACS é realizada pelo `genion`. O `genion`, ao processar as informações de topologia do sistema, realizará a substituição de moléculas de água pelos íons indicados pelo usuário. No entanto, o formato do arquivo de entrada do `genion` é do tipo `.tpr`. A criação do arquivo `.tpr` é realizado pelo `grompp`, através da linha de comando **`grompp -f ions.mdp -c EtAD_solv.gro -p topol.top -o ions.tpr`**. Note que linha de comando faz menção ao arquivo `"ions.mdp"`. Este se encontra na Fig. 3 e deverá ser fielmente transcrito para o Gedit, nomeado como `"ions"` e salvo com a extensão `.mdp` na pasta `"tutorial"` antes da execução da linha de comando.

```

; ions.mdp
integrator           = steep
emtol                = 1000.0
emstep              = 0.01
nsteps               = 50000
nstlist              = 1
ns_type              = grid
rlist                = 1.0
coulombtype          = PME
rcoulomb              = 1.0
rvdw                 = 1.0
pbc                  = xyz

```

Figura 3. Arquivo `"ions.mdp"`. O conteúdo desta figura deverá ser fielmente transcrito para o editor de texto Gedit. / **Figure 3.** File `"ions.mdp"`. The content of this figure should be faithfully transcribed into the Gedit text editor.

A geração do arquivo `.tpr` pelo `grompp` habilita a realização da etapa seguinte - a adição de íons para alcançar a neutralidade do sistema. As informações sobre a carga da proteína podem ser encontradas no arquivo de saída da ferramenta `pdb2gmx` gerado no início da análise. Alternativamente, a carga do sistema pode ser visualizada no terminal após a execução do `grompp` (procurar pela frase `"System has non-zero total charge: 2.000000"`). No caso da adenosina deaminase, a carga total sistema é de +2. A neutralização das cargas é feita através do comando **`genion -s ions.tpr -o EtAD_solv_ions.gro -p topol.top -pname NA -nname CL -nn 2`**. Por se tratar de uma ferramenta interativa, o `genion` retornará em resposta à linha de comando, uma tela com as opções de inserção dos íons. Neste tutorial, os íons deverão ser adicionados junto ao solvente (opção 13, `"SOL"`). No `genion`, o arquivo de entrada e de saída são marcados pelos comandos `-s` e `-o`, respectivamente. O comando `-p` indica o arquivo, no qual as moléculas de água serão substituídas por íons. Por último, os comandos `-pname` e `-nname` correspondem aos íons positivos e negativos a serem adicionados ao sistema, respectivamente.

Minimização da energia

O processo de solvatação gera um aumento da energia do sistema, além da possibilidade de gerar geometrias inapropriadas e impedimentos estéricos. Esses infortúnios são superados através da minimização da energia. Onde a energia das ligações é diminuída através de uma espécie de relaxamento molecular, tornando o sistema mais estável.

Para iniciar a minimização da energia, o usuário deverá transcrever as informações da figura 4 no Gedit e salvar o conteúdo como `"mini"` no formato `.mdp` e aloca-la na pasta `"tutorial"`. Após a criação do arquivo `mini.mdp`, a seguinte linha de

comando deverá ser inserida no terminal **"grompp -f mini.mdp -c EtAD_solv_ions.gro -p topol.top -o em.tpr"**. É importante que, após a execução do grompp, o usuário verifique no terminal a atualização do arquivo topol.top., indicada pela frase **"Back off! I just backed up topol.top to ./#topol.top.#"**. Após a finalização do grompp, o usuário deverá usar o comando **"mdrun -v -deffnm em"**, para iniciar a minimização de energia. Após a finalização, serão gerados quatro arquivos: i) em.log, apresenta o registro de atividade da minimização; ii) em.edr, arquivo binário de energia; iii) em.trr, arquivo binário de trajetória; iv) em.gro, estrutura minimizada. Para avaliar as informações contidas no arquivo em.edr, o usuário deverá gerar um gráfico (Fig. 5) com o uso do comando **"g_energy -f em.edr -o potential.xvg"** e, em seguida, escolher a opção "potential" (#10) acompanhado do número zero para finalizar. O gráfico poderá ser visualizado no software Xmgrace, que deverá ser instalado pelo usuário.

```

; mini.mdp
integrator          = steep
emtol               = 1000.0
emstep             = 0.01
nsteps             = 50000
nstlist            = 1
ns_type            = grid
rlist              = 1.0
coulombtype        = PME
rcoulomb           = 1.0
rvdw               = 1.0
pbc                = xyz

```

Figura 4. Arquivo "mini.mdp". O conteúdo desta figura deverá ser fielmente transcrito para o editor de texto Gedit. / **Figure 4.** File "mini.mdp". The content of this figure should be faithfully transcribed into the Gedit text editor.

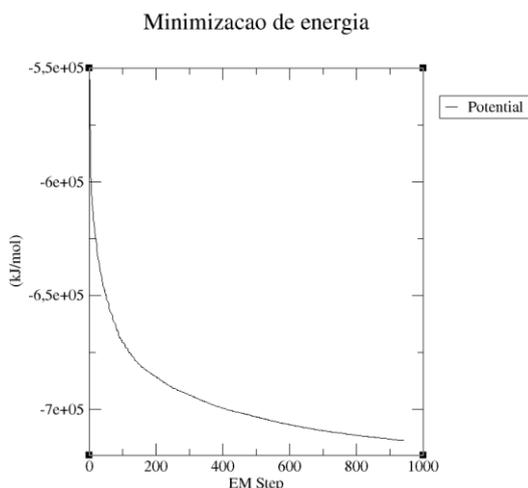


Figura 5. Potencial de energia de EtAD. O potencial de energia sempre deverá ser negativo. Para o caso de uma única proteína em água, a energia final deverá oscilar entre a ordem de 10^5 e 10^6 , dependendo do tamanho do sistema e do número de moléculas de água. / **Figure 5.** EtAD energy potential. The potential for energy should always be negative. In the case of a single protein in water, the final energy will oscillate between the order of 10^5 and 10^6 , depending on the system size and number of water molecules.

Equilíbrio

A minimização de energia é um passo importante na dinâmica molecular, especialmente acerca da orientação da geometria e orientação do solvente. No entanto, apenas a minimização de energia não habilita o início da simulação. Para a dinâmica molecular ocorrer sem problemas iniciais, é necessário equilibrar o solvente e os íons ao redor da proteína. Caso contrário, o sistema entrará em colapso durante a simulação.

Normalmente, a equalização (termo comumente usado em Biologia Computacional e Mecânica Molecular) é conduzida em duas fases. A primeira fase é realizada pelo conjunto canônico NVT (do inglês *constant Number of particles, Volume, and Temperature*). O tempo gasto no NVT é proporcional ao número de moléculas do sistema. No caso da EtAD, a conclusão do NVT é de aproximadamente 1 hora em computadores mais antigos e com baixa capacidade de processamento (por exemplo, Intel® Pentium® Dual Core). Para a primeira fase de equalização é necessário que o usuário transcreva as informações contidas na figura 6 para o Gedit, nomeie o arquivo como "NVT" e o salve com a extensão .mdp na pasta "tutorial". Para iniciar, o usuário deverá inserir a seguinte linha de comando no terminal **"grompp -f nvt.mdp -c em.gro -p topol.top -o nvt.tpr"**. Seguido do comando **"mdrun -deffnm nvt"**. A avaliação da progressão da temperatura pode ser realizada através do gráfico criado pelo comando **"g_energy -f nvt.edr"**. O usuário deverá selecionar a temperatura do sistema (opção 15) e sair (opção 0).

```

title              = OPLS EtAD NVT equilibration
define             = -DPOSRES
integrator         = md
nsteps            = 50000
dt                = 0.002
nstxout           = 100
nstvout           = 100
nstenergy         = 100
nstlog            = 100
continuation       = no
constraint_algorithm=lincs
constraints        = all.bonds
lincs_iter        = 1
lincs_order       = 4
ns_type           = grid
nstlist           = 5
rlist             = 1.0
rcoulomb          = 1.0
rvdw              = 1.0
coulombtype       = PME
pme_order         = 4
fourierspacing    = 0.16
tcoupl            = V-rescale
tc-grps           = Protein Non-Protein
tau_t             = 0.10.1
ref_t             = 300 300
pcoupl            = no
pbc               = xyz
DispCorr          = EnerPres
gen_vel           = yes
gen_temp          = 300
gen_seed          = -1

```

Figura 6. Arquivo "nvt.mdp". O conteúdo desta figura deverá ser fielmente transcrito para o editor de texto Gedit. / **Figure 6.** File "nvt.mdp". The content of this figure should be faithfully transcribed into the Gedit text editor.

Com o NVT a temperatura do sistema foi estabilizada, marcando o fim da primeira fase e habilitando o início da

segunda fase da equilibrção. A segunda fase é caracterizada pela equilíbrio da pressão. Essa etapa é conduzida pelo conjunto NPT (do inglês *Number of particles, Pressure, and Temperature*) sendo todos os elementos do acrônimo, constantes. O tempo para conclusão do NPT é semelhante ao gasto pelo NVT. Antes de iniciar a segunda fase de equilibrção é necessário que o usuário transcreva as informações contidas na figura 7 para o Gedit, nomeie o arquivo como "NPT" e o salve com a extensão .mdp na pasta "tutorial". A equilibrção da pressão tem início após a inserção da linha de comando "**grompp -f npt.mdp -c nvt.gro -t nvt.cpt -p topol.top -o npt.tpr**", seguido do comando "**mdrun -deffnm npt**". A visualização da progressão da pressão da EtAD pode ser feita após o término do NPT pelo comando "**g_energy -f npt.edr -o pressure.xvg**", seguida das opções 16 (pressão do sistema) e 0 (comando de saída).

```

title = OPLS EtAD NPT equilibration
define = -DPOSRES
integrator = md
nsteps = 50000
dt = 0.002
nstxout = 100
nstvout = 100
nstenergy = 100
nstlog = 100
continuation = yes
constraint_algorithm=lincs
constraints = all.bonds
lincs_iter = 1
lincs_order = 4
ns_type = grid
nstlist = 5
rlist = 1.0
rcoulomb = 1.0
rvdw = 1.0
coulombtype = PME
pme_order = 4
fourierspacing = 0.16
tcoupl = V-rescale
tc-grps = Protein Non-Protein
tau_t = 0.10.1
ref_t = 300 300
pcoupl = Parrinello-Rahman
pcoupltype = isotropic
tau_p = 2.0
ref_p = 1.0
compressibility = 4.5e-5
refcoord_scaling = com
pbc = xyz
DispCorr = EnerPres
gen_vel = no

```

Figura 7. Arquivo "npt.mdp". O conteúdo desta figura deverá ser fielmente transcrito para o editor de texto Gedit. / **Figure 7.** File "npt.mdp". The content of this figure should be faithfully transcribed into the Gedit text editor.

Dinâmica Molecular

Todos os processos realizados acima são necessários para a realização da simulação. Após completar as fases de

equilibrção, o sistema se apresenta com a temperatura e pressão desejadas. Conseqüentemente, o sistema está apto para liberar as posições de restrição e iniciar a dinâmica molecular para a coleta de dados. Neste caso será calculado 1 nanosegundo (ns) de simulação. Trata-se de um processo demorado, aproximadamente 13 horas em computadores antigos (Intel® Pentium® Dual Core). Antes de iniciar a dinâmica molecular é necessário que o usuário transcreva as informações contidas na figura 8 para o Gedit, nomeie o arquivo como "md" e o salve com a extensão .mdp na pasta "tutorial". O início da simulação é marcado pelo comando "**grompp -f md.mdp -c npt.gro -t npt.cpt -p topol.top -o md_0_1.tpr**", seguido de "**mdrun -deffnm md_0_1**". A versão do GROMACS utilizada neste estudo possui aceleração via GPU (*Graphics Processing Unit*). Caso o usuário possua uma GPU disponível, poderá ser usado o seguinte comando, em alternativa ao mdrun anterior, "**mdrun -deffnm md_0_1 -nb gpu**".

```

title = OPLS EtAD MD
integrator = md
nsteps = 500000
dt = 0.002
nstxout = 1000
nstvout = 1000
nstenergy = 1000
nstlog = 1000
continuation = yes
constraint_algorithm=lincs
constraints = all.bonds
lincs_iter = 1
lincs_order = 4
ns_type = grid
nstlist = 5
rlist = 1.0
rcoulomb = 1.0
rvdw = 1.0
coulombtype = PME
pme_order = 4
fourierspacing = 0.16
tcoupl = V-rescale
tc-grps = Protein Non-Protein
tau_t = 0.1 0.1
ref_t = 300 300
pcoupl = Parrinello-Rahman
pcoupltype = isotropic
tau_p = 2.0
ref_p = 1.0
compressibility = 4.5e-5
pbc = xyz
DispCorr = EnerPres
gen_vel = no

```

Figura 8. Arquivo "md.mdp". O conteúdo desta figura deverá ser fielmente transcrito para o editor de texto Gedit. / **Figure 8.** File "mdf.mdp". The content of this figure should be faithfully transcribed into the Gedit text editor.

Ao fim da dinâmica molecular é necessário realizar o pós-processamento dos dados, através da linha de comando "**trjconv -s md_0_1.tpr -f md_0_1.xtc -o mdnoPBC.xtc -pbc mol -ur compact**", seguido da escolha da opção 0 (*system*) para o arquivo

de saída. A partir desse ponto, várias análises poderão ser realizadas, como por exemplo, RMSD (*Root Mean Deviation Square*), para avaliar sua estabilidade da proteína e raio de giro, útil na mensuração da qualidade do enovelamento proteico. A visualização do RMSD (Fig. 9) pode ser feita através do comando “**g_rms -s md_0_1.tpr -f mdnoPBC.xtc -o rmsd.xvg -tu ns**”, seguido da escolha do campo 4 (*backbone*) para ambas solicitações. Os dados demonstram que a EtAD é estável durante os 1.000 picosegundos (ps) de simulação, tendo em vista que a oscilação do RMSD não ultrapassou 0,1 nm (1Å) do valor inicial. Já a visualização do raio de giro pode ser realizada pelo comando “**g_gyrate -s md_0_1.tpr -f mdnoPBC.xtc -o gyrate.xvg**” (Fig. 10). O raio de giro da EtAD permaneceu estável, revelando o alto grau de compactidade proteica, reflexo do enovelamento estável da proteína.

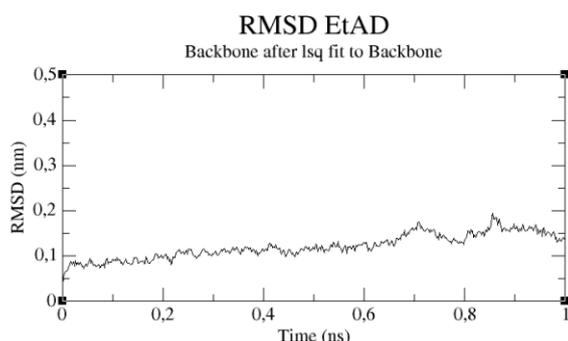


Figura 9. Flutuação do RMSD da EtAD em função do tempo. Os níveis do RMSD flutuam em torno de 0,1 e 0,15 nm, indicando o alto grau de estabilidade estrutural da EtAD. / **Figure 9.** fluctuation RMSD of EtAD a function of time. The RMSD levels fluctuate around 0.1 nm and 0.15, indicating the high degree of structural stability etad.

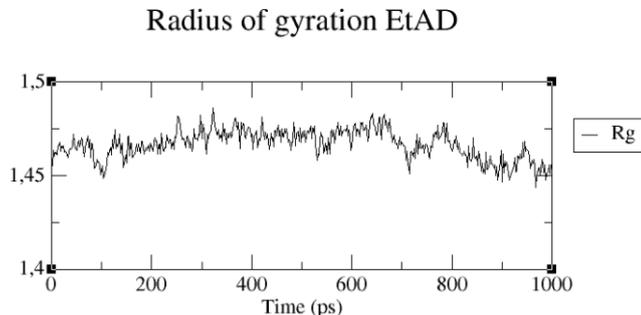


Figura 10. Raio de giro da EtAD. O raio de giro permanece estável durante a simulação, indicando a alta qualidade do enovelamento proteico. / **Figure 10.** Turning radius of EtAD. The turning radius remains stable during the simulation, indicating the high quality of protein folding.

4. Conclusão

Este trabalho trata da primeira simulação envolvendo uma estrutura proteica de *E. tarda*. Os resultados gerados na simulação indicam que a EtDA é uma estrutura altamente estável, como demonstrado pelo RMSD e raio de giro. Esses dados poderão conduzir, em estudos futuros, ao melhor entendimento dos processos bioquímicos e fisiológicos de *E. tarda*, bem como, auxiliar o desenvolvimento de estratégias para o controle e/ou eliminação desta bactéria. Além disso, este trabalho proporciona ao usuário um contato inicial e facilitado com a dinâmica molecular, sobretudo, no que se refere à execução da simulação, tendo em vista que o GROMACS não possui interface gráfica.

5. Referências Bibliográficas

CHENG, C. J.; DAGGETT, V. Molecular Dynamics Simulations Capture the Misfolding of the Bovine Prion Protein at Acidic pH. **Biomolecules**, v.

4, n. 1, p. 181-201, 2014.

DAURA, X. et al. On the sensitivity of MD trajectories to changes in water-protein interaction parameters: the potato carboxypeptidase inhibitor in water as a test case for the GROMOS force field. **Proteins**, v. 25, n. 1, p. 89-103, May 1996.

EISENBERG, D. et al. VERIFY3D: assessment of protein models with three-dimensional profiles. **Methods Enzymol**, v. 277, p. 396-404, 1997.

ELLEBY, B. et al. Enhancement of catalytic efficiency by the combination of site-specific mutations in a carbonic anhydrase-related protein. **European Journal of Biochemistry**, v. 267, n. 19, p. 5908-15, Oct 2000.

GUZMAN, I.; GRUEBELE, M. Protein Folding Dynamics in the Cell. **Journal of Physical Chemistry B**, Jun 17 2014.

HU, Y. H. et al. *Edwardsiella tarda* Hfq: impact on host infection and global protein expression. **Veterinary Research**, v. 45, p. 23, 2014.

HUB, J. S.; DE GROOT, B. L. Detection of functional modes in protein dynamics. **PLOS Computational Biology**, v. 5, n. 8, p. e1000480, 2009.

HUMPHREY, W. et al. VMD: visual molecular dynamics. **Journal of Molecular Graphics**, v. 14, n. 1, p. 33-8, 27-8, Feb 1996.

JANDA, J. M.; ABBOTT, S. L. Infections associated with the genus *Edwardsiella*: the role of *Edwardsiella tarda* in human disease. **Clinical Infectious Diseases**, v. 17, n. 4, p. 742-8, Oct 1993.

JANSSON, H. et al. Dynamics of a protein and its surrounding environment: a quasielastic neutron scattering study of myoglobin in water and glycerol mixtures. **Journal of Chemical Physics**, v. 130, n. 20, p. 205101, May 28 2009.

JARURATANASIRIKUL, S.; KALNAUWAKUL, S. *Edwardsiella tarda*: a causative agent in human infections. **Southeast Asian J Trop Med Public Health**, v. 22, n. 1, p. 30-4, Mar 1991.

KAUSHIK, A. et al. *In silico* characterization and molecular dynamics simulation of Pfcyc-1, a cyclin homolog of *Plasmodium falciparum*. **Journal of Biomolecular Structure and Dynamics**, v. 32, n. 10, p. 1624-33, 2014.

KUMAR, H. et al. Design, Synthesis, Physicochemical Studies, Solvation, and DNA Damage of Quinoline-Appended Chalcone Derivative: Comprehensive Spectroscopic Approach toward Drug Discovery. **Journal of Physical Chemistry B**, Jun 25 2014.

LANGE, O. F. et al. Flooding in GROMACS: accelerated barrier crossings in molecular dynamics. **Journal of Computational Chemistry**, v. 27, n. 14, p. 1693-702, Nov 15 2006.

LASKOWSKI, R. A. et al. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. **Journal of Biomolecular NMR**, v. 8, n. 4, p. 477-86, Dec 1996.

LUO, M.; SCHRAMM, V. L. Transition state structure of *E. coli* tRNA-specific adenosine deaminase. **Journal of the American Chemical Society**, v. 130, n. 8, p. 2649-55, Feb 27 2008.

MACCALLUM, J. L.; TIELEMAN, D. P. Calculation of the water-cyclohexane transfer free energies of neutral amino acid side-chain analogs using the OPLS all-atom force field. **Journal of Computational Chemistry**, v. 24, n. 15, p. 1930-5, 2003.

NADEAU, O. W.; CARLSON, G. M. A review of methods used for identifying structural changes in a large protein complex. **Methods in Molecular Biology**, v. 796, p. 117-32, 2012.

NELSON, J. J. et al. Extraintestinal manifestations of *Edwardsiella tarda* infection: a 10-year retrospective review. **Journal of the Louisiana State Medical Society**, v. 161, n. 2, p. 103-6, 2009.

PETERSEN, E. F. et al. UCSF Chimera--a visualization system for exploratory research and analysis. **Journal of Computational Chemistry**, v. 25, n. 13, p. 1605-12, 2004.

PHILLIPS, J. C. et al. Scalable molecular dynamics with NAMD. **Journal of Computational Chemistry**, v. 26, n. 16, p. 1781-802, 2005.

ROTERMAN, I. K. et al. A comparison of the CHARMM, AMBER and ECEPP potentials for peptides. I. Conformational predictions for the tandemly repeated peptide (Asn-Ala-Asn-Pro)₉. **Journal of Biomolecular Structure and Dynamics**, v. 7, n. 3, p. 391-419, 1989.

SALI, A. et al. Evaluation of comparative protein modeling by MODELLER. **Proteins**, v. 23, n. 3, p. 318-26, Nov 1995.

VAN DER SPOEL, D. et al. GROMACS: fast, flexible, and free. **Journal of Computational Chemistry**, v. 26, n. 16, p. 1701-18, 2005.

-
- TEKEDAR, H. C. et al. Complete Genome Sequence of Channel Catfish Gastrointestinal Septicemia Isolate *Edwardsiella tarda* C07-087. **Genome Announc**, v. 1, n. 6, 2013.
- XU, Z. et al. Modifying the OPLS-AA force field to improve hydration free energies for several amino acid side chains using new atomic charges and an off-plane charge model for aromatic residues. **Journal of Computational Chemistry**, v. 28, n. 3, p. 689-97, 2007.
- WANG, C. et al. Identification of plasma-responsive outer membrane proteins and their vaccine potential in *Edwardsiella tarda* using proteomic approach. **Journal of Proteomics**, v. 75, n. 4, p. 1263-75, Feb 2 2012.
- WANG, J. et al. Development and testing of a general amber force field. **Journal of Computational Chemistry**, v. 25, n. 9, p. 1157-74, 2004.
- WIEDERSTEIN, M.; SIPPL, M. J. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. **Nucleic Acids Research**, v. 35, n. Web Server issue, p. W407-10, 2007.
- ZHENG, Y. et al. Different approaches to expressing *Edwardsiella tarda* antigen GAPDH in attenuated *Vibrio anguillarum* for multivalent fish vaccines. **Journal of Fish Diseases**, v. 35, n. 8, p. 569-77, 2012.