

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

Tatiane Fernanda de Souza Silva

**Modelagem molecular: predição de estruturas
de peptídeos pelo método *ab initio***

Uberlândia, Brasil

2019

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

Tatiane Fernanda de Souza Silva

**Modelagem molecular: predição de estruturas de
peptídeos pelo método *ab initio***

Trabalho de conclusão de curso apresentado
à Faculdade de Computação da Universidade
Federal de Uberlândia, Minas Gerais, como
requisito exigido parcial à obtenção do grau
de Bacharel em Ciência da Computação.

Orientador: Prof. Dr. Anderson Rodrigues dos Santos

Universidade Federal de Uberlândia – UFU

Faculdade de Computação

Bacharelado em Ciência da Computação

Uberlândia, Brasil

2019

Tatiane Fernanda de Souza Silva

Modelagem molecular: predição de estruturas de peptídeos pelo método *ab initio*

Trabalho de conclusão de curso apresentado à Faculdade de Computação da Universidade Federal de Uberlândia, Minas Gerais, como requisito exigido parcial à obtenção do grau de Bacharel em Ciência da Computação.

Trabalho aprovado. Uberlândia, Brasil, 10 de julho de 2019:

Prof. Dr. Anderson Rodrigues dos Santos
Orientador

Prof. Dr. Eduardo de Faria Franca
Instituto de Química - UFU

Dino Rogério Coinete Franklin
FACOM - UFU

Uberlândia, Brasil
2019

*Dedico este trabalho aos meu pais,
Marquilandes e Enilza,
que me apoiaram em todos os momentos da minha vida.
E ao meu namorado,
Rodrigo,
que me apresentou o mundo da computação.
Amo vocês.*

Agradecimentos

Quero agradecer Deus e a todos aqueles que me apoiaram de forma direta e indiretamente a concluir este trabalho, dentre os quais são:

Meu orientador, que me instruiu e me apoiou em todos os momentos de dúvida, acompanhando sempre o meu progresso.

Meus professores do curso de Ciência da Computação que, através dos seus ensinamentos, permitiram que eu pudesse concluir este trabalho.

Minha família, meu namorado e amigos que tiveram paciência comigo nos momentos difíceis, me apoiando e me dando força.

“Na vida, não existe nada a temer, mas a entender.”
(Marie Curie)

Resumo

Os programas que são considerados o estado da arte conseguem fazer predições de estruturas tridimensionais de até 56 aminoácidos utilizando conceitos de campos de força como: atração eletrostática, van der Waals, restrições de ângulos de torção de carbonos cadeias alfa e beta de aminoácidos e comprimento da ligação. Estudamos uma forma alternativa para fazer a predição da estrutura tridimensional aproximada de peptídeos e desenvolvemos um programa que faz essa predição com uma sequência aminoácidos utilizando Algoritmo Genético(AG). Comparamos os resultados obtidos através do programa com estruturas conhecidas, extraídas da base de dados PubChem, comparando a energia eletrônica total e monitorando o tempo de execução de acordo com que aumentamos o tamanho da molécula. Obtivemos resultados bem próximos em termos de energia, mas o tempo de execução aumenta significativamente da entrada de dois para três aminoácidos. Isso já era esperado, pois o problema da Modelagem Molecular pertence à classe NP-Completo e o tempo para resolver o mesmo aumenta exponencialmente à medida que aumentamos o tamanho da molécula.

Palavras-chave: Modelagem Molecular. Peptídeos. Método *ab initio*. Algoritmos Genéticos. Química Computacional.

Abstract

Programs that are considered the state of art can predict three-dimensional structures of up to 56 amino acids using force field concepts such as electrostatic attraction, van der Waals, torsion angle constraints of carbon alpha and beta chains of amino acids and bond length. We studied an alternative way to predict the approximate three-dimensional structure of peptides and developed a program that makes this prediction with an amino acid sequence using Genetic Algorithm (GA). We compared the results obtained through the program with known structures, extracted from the PubChem database, comparing the total electronic energy and monitoring the execution time as we increase the size of the molecule. We get very close results in terms of energy, but the runtime increases significantly from the input of two to three amino acids. This is already expected because the problem of Molecular Modeling belongs to the NP-Complete class and the time to solve it increases exponentially as we increase the size of the molecule.

Key-words: Molecular Modeling. Peptides. *Ab initio* method. Genetic Algorithms. Computational Chemistry.

Lista de ilustrações

Figura 1 – Conformações que podem ser geradas de um mesmo peptídeo variando os ângulos diedro ϕ e ψ . Fonte: Adaptada de (NELSON; COX, 2014)	14
Figura 2 – Variações de energia das rotações que surgem em torno da ligação C2 e C3 do butano. Fonte: Adaptada de (SOLOMONS; FRYHLE, 2000)	18
Figura 3 – Ilustração 3D da molécula do butano com os respectivos carbonos numerados	19
Figura 4 – Isomeria do but-2-eno. Fonte Adaptada de (ISOMERIA..., 2018)	19
Figura 5 – Estrutura geral de um aminoácido. Fonte: Adaptada de (AMINOÁCIDOS, 2017)	20
Figura 6 – Formação de uma ligação peptídica por condensação. Fonte: Extraída de (NELSON; COX, 2014)	21
Figura 7 – Níveis de estrutura. Fonte: Extraída de (NELSON; COX, 2014)	21
Figura 8 – Exemplo de dois aminoácidos formando um dipeptídeo	25
Figura 9 – Exemplo 1 de três aminoácidos formando um tripeptídeo	25
Figura 10 – Exemplo 2 de três aminoácidos formando um tripeptídeo	26
Figura 11 – Exemplo de rotações das ligações peptídicas utilizadas no Algoritmo Genético(AG).	26
Figura 12 – Fluxograma do Algoritmo Genético(AG) implementado no programa.	28
Figura 13 – Projeção 3D da molécula Ala-Gly.	31
Figura 14 – Projeção 3D da molécula Ala-Gly-Ser.	34
Figura 15 – Projeção 3D da molécula Ala-Gly-Ser-Glu.	37
Figura 16 – Gráfico que mostra o tempo de execução para cada entrada do programa.	40

Lista de tabelas

Tabela 1 – Funções de polarização para vários átomos(cc-pVDZ). Fonte: adaptado de (GAUSSIAN... , 2017).	17
Tabela 2 – Simbologia dos aminoácidos. Fonte: extraído de (AMINOÁCIDOS , 2017)	20
Tabela 3 – Resultante do programa.	32
Tabela 4 – Base PubChem.	32
Tabela 5 – Matriz das coordenadas da molécula Ala-Gly	32
Tabela 6 – Ligações da molécula Ala-Gly	33
Tabela 7 – Resultante do programa.	35
Tabela 8 – Base PubChem.	35
Tabela 9 – Matriz normalizada das coordenadas da molécula Ala-Gly-Ser	35
Tabela 10 – Ligações da molécula Ala-Gly-Ser	36
Tabela 11 – Resultante do programa.	38
Tabela 12 – Base PubChem.	38
Tabela 13 – Matriz normalizada das coordenadas da molécula Ala-Gly-Ser-Glu	38
Tabela 14 – Ligações da molécula Ala-Gly-Ser-Glu	39

Lista de abreviaturas e siglas

3D	Tridimensional
A	Alanina
AEs	Algoritmos Evolutivos
AGs	Algoritmos Genéticos
AHP	Analytical Hierarchy Process
C	Carbono
CASP	Critical Assessment of Structural Prediction
CNTP	Condições Normais de Temperatura e Pressão
COOH	Ácido Carboxílico
G	Glicina
HF	Hartree-Fock
HMM	Hidden Markov Model
IA	Inteligência Artificial
IUPAC	International Union of Pure and Applied Chemistry
MM	Modelagem Molecular
N	Nitrogênio
SA	Structural Alphabet
SCF	Self-Consistent Field
UML	Unified Modeling Language

Lista de símbolos

\AA	Ângström
C_α	Carbono alfa
ϕ, ψ	Ângulos diedro ou de torção

Sumário

1	INTRODUÇÃO	14
1.1	Objetivos	14
1.1.1	Objetivo Geral	14
1.1.2	Objetivo Específico	15
1.2	Justificativa	15
2	FUNDAMENTAÇÃO TEÓRICA	16
2.1	Modelagem Molecular	16
2.1.1	Método de Hartree-Fock	16
2.1.2	Método <i>ab initio</i>	17
2.1.3	Análise Conformacional	17
2.2	Isomerismo	19
2.3	Aminoácidos	19
2.4	Peptídeos	20
2.4.1	Classificação estrutural	21
2.5	Algoritmos Evolutivos	22
3	DESENVOLVIMENTO	23
3.1	Sistema Operacional e Hardware	23
3.2	Bibliotecas de Terceiros	23
3.2.1	Psi4NumPy	23
3.2.2	SciPy	23
3.3	Estratégia para a modelagem	24
3.4	Implementação	26
3.5	Resultados	31
4	CONCLUSÃO	41
	REFERÊNCIAS	42
	ANEXOS	44
	ANEXO A – INSTALAÇÃO DAS BIBLIOTECAS NECESSÁRIAS PARA A UTILIZAÇÃO DO PROGRAMA	45
A.1	Psi4NumPy	45
A.2	SciPy	45

	ANEXO B – UTILIZAÇÃO DA BIBLIOTECA PSI4	46
	ANEXO C – SAÍDAS DO PROGRAMA IMPLEMENTADO NESSE	
	TRABALHO	47
C.1	Ala-Gly	47
C.2	Ala-Gly-Ser	57
C.3	Ala-Gly-Ser-Glu	69

1 Introdução

A pesquisa em busca de métodos computacionais mais eficientes para problemas complexos é de grande importância, como o problema da predição da estrutura tridimensional(3D) de peptídeos. A busca pela solução ótima, no caso da predição das estruturas 3D dos peptídeos, é algo que pode ser difícil de encontrar, pois esses problemas são considerados NP-completo(BERGER; LEIGHTON, 1998), onde os métodos de otimização tradicionais não costumam ter um bom resultado.

Os progressos na área de bioinformática que estuda a predição da estrutura 3D das proteínas são monitorados com um teste bianual chamado de competição CASP (Critical Assessment of Structural Prediction, ou Avaliação Crítica de Predição Estrutural)(NELSON; COX, 2014). Um dos métodos utilizados para essa predição é o método *ab initio*. Eles são soluções aproximadas baseadas na equação de Schrödinger em que não utilizam dados experimentais(JENSEN, 2007). Uma combinação de métodos podem gerar resultados ainda mais eficientes, com isso, podemos associar o método *ab initio* com o uso de algoritmos genéticos(AGs), uma solução interessante, pois são baseados no princípio da Evolução de Darwin.

Os algoritmos genéticos podem ser utilizados na etapa de busca das conformações de menor energia dos peptídeos. Depois de gerado uma estrutura aproximada pelo método *ab initio*, variações nos ângulos diedro ou de torção(ϕ, ψ) geram várias conformações de uma mesma molécula.

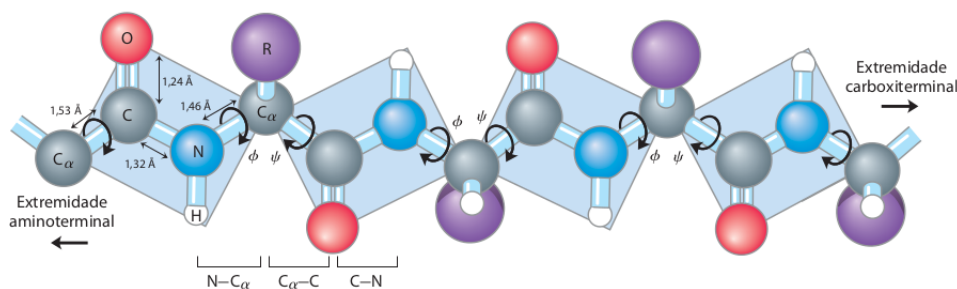


Figura 1 – Conformações que podem ser geradas de um mesmo peptídeo variando os ângulos diedro ϕ e ψ . Fonte: Adaptada de (NELSON; COX, 2014)

1.1 Objetivos

1.1.1 Objetivo Geral

Desenvolver um programa que faça a predição da estrutura tridimensional aproximada de peptídeos e estudar métodos alternativos para resolver o problema da predição.

1.1.2 Objetivo Específico

Os objetivos específicos desse trabalho são:

- Estudo da Modelagem Molecular.
- Estudo do método *ab initio* para a predição de peptídeos.
- Utilização de Algoritmos Genéticos para obter a estrutura 3D de peptídeos com conformações de menor energia.
- Comparar resultados com estruturas já conhecidas na literatura.

1.2 Justificativa

Atualmente, existem muitos trabalhos que investigam as estruturas das moléculas, utilizando diferentes metodologias, para ser aplicada na predição das estruturas 3D de peptídeos e proteínas.

Por exemplo, o trabalho que criou a ferramenta PEP-FOLD1 ([MAUPETIT; DERREUMAUX; TUFFERY, 2009](#)) utilizou o método *de novo* com auxílio de um Hidden Markov Model(HMM) para predição de uma estrutura inicial e uma modificação do algoritmo guloso para obter as estruturas 3D aproximadas que serão refinadas por uma simulação Monte Carlo de 300 mil passos.

Em PEP-FOLD2 ([THÉVENET et al., 2012](#)), o sistema foi melhorado para tratar peptídeos cíclicos, lineares e com ligações dissulfeto.

Já em Sun ([SUN, 1995](#)), a busca pela melhor conformação de peptídeos e proteínas foi feita usando um método baseado em algoritmos genéticos(AGs).

Na Otimização Combinatória e Teoria da Complexidade, o problema de predição de cadeia de aminoácidos é classificado como um problema NP-Completo, devido às limitações de predições experimentais das estruturas 3D dos peptídeos e a grande necessidade de conhecer essas estruturas([CRESCENZI et al., 1998](#)). Isso significa que o tempo para o cálculo depende do tamanho da cadeia de aminoácidos, e ele cresce de forma exponencial impedindo a predição de estruturas complexas.

Assim, a principal justificativa deste trabalho é estudar alternativas para a predições das estruturas dos peptídeos e implementar um protótipo de programa com uma dessas alternativas para que seja possível analisar seus resultados.

2 Fundamentação Teórica

Neste capítulo é apresentado os conceitos básicos necessários para compreensão deste trabalho.

2.1 Modelagem Molecular

A Modelagem Molecular, segundo a IUPAC (International Union of Pure and Applied Chemistry), consiste na investigação das estruturas e das propriedades moleculares pelo uso de química computacional, visando fornecer uma representação tridimensional, sob dadas circunstâncias(CARVALHO et al., 2003)(MODELAGEM. . . , 2016).

2.1.1 Método de Hartree-Fock

O método Hartree-Fock(HF) foi criado pelo matemático Douglas Hartree e o físico Vladimir Fock, tendo como objetivo produzir Orbitais Moleculares otimizados(SMITH et al., 2018). Ele é utilizado como base para calcular a energia do estado fundamental de átomos e moléculas.

Infelizmente, o método HF possui várias dificuldades em sua aplicação. Por exemplo, é necessário um conjunto completo de funções de onda de partículas únicas para calcular o potencial não-local de um único elétron. E também é necessário seguir um procedimento bastante complexo para incluir as correções de correlação que estão além da estrutura de HF (AMUSIA; MSEZANE; SHAGINYAN, 2003). Com isso, o cálculo da energia de sistemas que possuem muitos elétrons, como moléculas com vários átomos, se tornam muito complexos.

A função de Hartree-Fock é dada por,

$$E_{HF} = \langle \psi_0 | \hat{H} | \psi_0 \rangle = \sum_i \langle i | \hat{h} | i \rangle + \frac{1}{2} \sum_{ij} [ii|jj] - [ij|ji] \quad (2.1)$$

onde \hat{H} é o Hamiltoniano molecular(SMITH et al., 2018).

A teoria do Campo Autoconsistente(*self-consistent field*, SCF) forma a base da química quântica *ab initio*(SMITH et al., 2018), conduzindo as equações de HF, também conhecido como HF-SCF. Assim, procuramos por uma solução convergente que calcule iterativamente o potencial elétrico produzido pela distribuição de cargas dada a densidade de probabilidade obtida dos orbitais espaciais.

A maioria dos métodos HF-SCF requer que um conjunto de base seja especificado, nesse trabalho utilizamos o conjunto de funções base duplo zeta **cc-pvdz**, que possui o dobro de funções mínimas para descrever cada orbital. Esse conjunto base inclui funções de polarização por definição.

Átomos	Orbitais
H	2s,1p
He	2s,1p
Li-Be	3s,2p,1d
B-Ne	3s,2p,1d
Na-Ar	4s,3p,1d
Ca	5s,4p,2d
Sc-Zn	6s,5p,3d, 1f
Ga-Kr	5s,4p,2d

Tabela 1 – Funções de polarização para vários átomos(**cc-pVDZ**). Fonte: adaptado de (GAUSSIAN..., 2017).

2.1.2 Método *ab initio*

Os métodos *ab initio*, são descritos como métodos da química computacional baseados na química quântica(LEVINE, 1991)(MÉTODOS..., 2017). Eles são derivados diretamente de princípios teóricos, como a equação de Schrödinger, sem inclusão de dados experimentais.

Esse método, de fato, pode ser visto como um método aproximado da mecânica quântica. As aproximações feitas geralmente são aproximações matemáticas, como usar uma forma funcional mais simples para uma função ou obter uma solução aproximada para uma equação diferencial(RAMACHANDRAN; DEEPA; NAMBOORI, 2008).

Uma modelagem *ab initio* bem-sucedida depende de três fatores(RIGDEN, 2017):

- Uma função de energia precisa, equação 2.1, com a qual a estrutura nativa de uma molécula corresponde ao estado mais termodinamicamente estável, em comparação com todas as possíveis estruturas;
- Um método de busca eficiente que pode identificar rapidamente os estados de baixa energia através da busca conformacional;
- Seleção de modelos nativos de um conjunto de estruturas;

2.1.3 Análise Conformacional

Uma mesma molécula pode assumir mais de uma estrutura 3D, mas nem todas são estáveis em condições normais de temperatura e pressão(CNTP). Essas formas estruturais,

resultados das rotações dos grupos funcionais em torno de ligações simples, são chamadas **conformações** da molécula(SOLOMONS; FRYHLE, 2000). Cada conformação possui uma energia potencial associada.

A **análise conformacional** é o estudo da variação de energia que a molécula sofre com os grupos funcionais girando sobre uma ligação simples(SOLOMONS; FRYHLE, 2000).

Para exemplificar isso, podemos observar a análise conformacional do butano na figura 2:

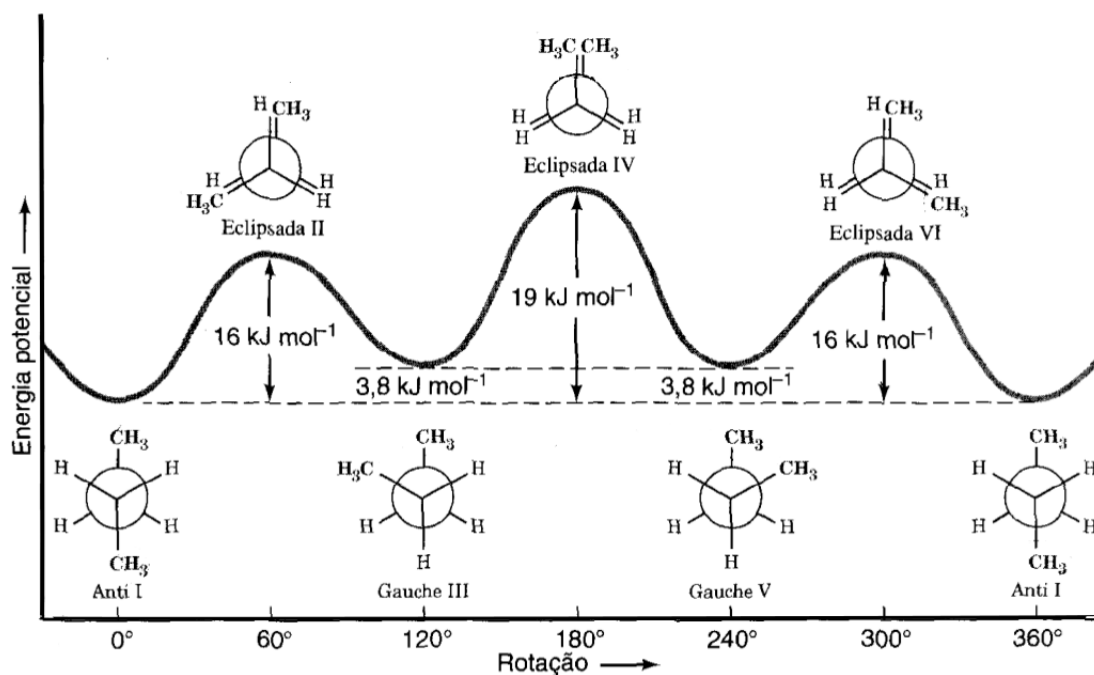


Figura 2 – Variações de energia das rotações que surgem em torno da ligação C2 e C3 do butano. Fonte: Adaptada de (SOLOMONS; FRYHLE, 2000)

Ao fazer rotações em torno da ligação C2 e C3 do butano, a torção entre os átomos é alterada. Com isso, as forças intermoleculares mudam precisando de uma maior ou menor energia para manter a molécula naquela conformação.

A figura 3 representa o butano em sua forma 3D, onde as esferas de cor cinza são os átomos de carbono e as de cor branca são os átomos de hidrogênio. Os carbonos foram numerados de acordo com as regras da IUPAC.

Assim, podemos observar na figura 2, que a conformação do tipo **anti** é a mais estável, pois é a que necessita de menor energia para manter a geometria no espaço 3D. Isso porque os grupos metila(CH₃) estão em posições opostas, então não existe tensão torsional nessa conformação.

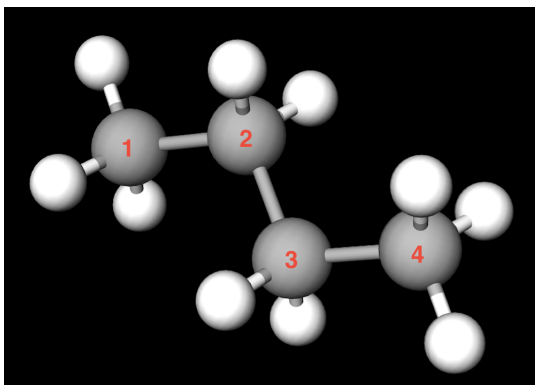
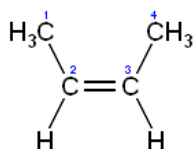


Figura 3 – Ilustração 3D da molécula do butano com os respectivos carbonos numerados

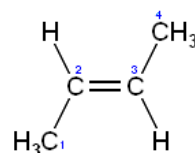
2.2 Isomerismo

O isomerismo ou isomeria é um fenômeno no qual dois ou mais compostos químicos diferentes denominados isômeros apresentam a mesma fórmula molecular (mesmo conjunto de átomos) e diferentes fórmulas estruturais (diferente arranjo entre os átomos) (GARCIA CLEVERSON FERNANDO; FERREIRA LUCAS, 2015).

Na **isomeria cis-trans** os isômeros possuem a mesma fórmula molecular e também a mesma fórmula estrutural plana, o que diferencia é apenas as fórmulas estruturais geométrica. Os isômeros também apresentam propriedades físicas e químicas diferentes por apresentar estruturas espaciais diferentes.



(a) *cis*-but-2-eno(IUPAC).



(b) *trans*-but-2-eno(IUPAC).

Figura 4 – Isomeria do but-2-eno. Fonte Adaptada de (ISOMERIA..., 2018).

2.3 Aminoácidos

Aminoácidos são moléculas que contêm em sua estrutura um grupo amina (NH_2) e um grupo carboxila ($COOH$), ambos ligados ao carbono α (C_α), que é o carbono adjacente ao grupo funcional. O C_α também é ligado a um hidrogênio (H) e a uma cadeia lateral R (radical).

Eles se diferem uns dos outros em suas cadeias laterais, que variam em estrutura, tamanho e carga elétrica, e que influenciam a solubilidade dos aminoácidos em água (NELSON; COX, 2014).

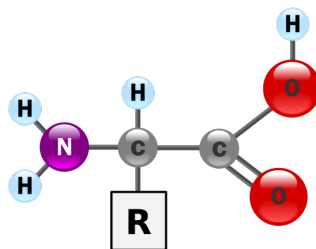


Figura 5 – Estrutura geral de um aminoácido. Fonte: Adaptada de ([AMINOÁCIDOS, 2017](#))

Nome	Símbolo	Abreviação
Glicina ou Glicocola	Gly, Gli	G
Alanina	Ala	A
Leucina	Leu	L
Valina	Val	V
Isoleucina	Ile	I
Prolina	Pro	P
Fenilalanina	Phe ou Fen	F
Serina	Ser	S
Treonina	Thr, Tre	T
Cisteína	Cys, Cis	C
Tirosina	Tyr, Tir	Y
Asparagina	Asn	N
Glutamina	Gln	Q
Aspartato ou Ácido aspártico	Asp	D
Glutamato ou Ácido glutâmico	Glu	E
Arginina	Arg	R
Lisina	Lys, Lis	K
Histidina	His	H
Triptofano	Trp, Tri	W
Metionina	Met	M

Tabela 2 – Simbologia dos aminoácidos. Fonte: extraído de ([AMINOÁCIDOS, 2017](#))

A ligação de vários aminoácidos gera uma macromolécula chamada de proteína. Elas são construídas a partir do mesmo conjunto onipresente de 20 aminoácidos(Tabela 2). Cada um desses aminoácidos possui uma cadeia lateral com propriedades químicas características, esse grupo de 20 moléculas precursoras pode ser considerado o alfabeto no qual a linguagem da estrutura proteica é lida.([NELSON; COX, 2014](#))

2.4 Peptídeos

Os peptídeos são compostos formados por dois ou mais aminoácidos. Para produzir um **dipeptídeo**, duas moléculas de aminoácidos podem ser ligadas de modo covalente

por meio de uma ligação amida substituída, denominada ligação peptídica. Essa ligação é formada pela remoção de elementos de água (desidratação) do grupo α -carboxila de um aminoácido e do grupo α -amino do outro (NELSON; COX, 2014).

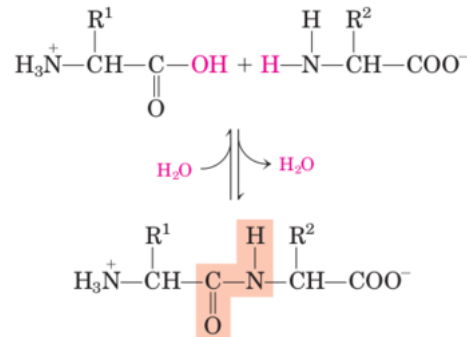


Figura 6 – Formação de uma ligação peptídica por condensação. Fonte: Extraída de (NELSON; COX, 2014)

A estrutura de um peptídeo formada por ligações de alguns aminoácidos é chamada de **oligopeptídeo**, já a formada pela ligação de muitos aminoácidos é chamada de **polipeptídeo**.

2.4.1 Classificação estrutural

Dependendo da quantidade de aminoácidos e as interações entres eles, podemos classificar os peptídeos em quatro estruturas, figura 7.

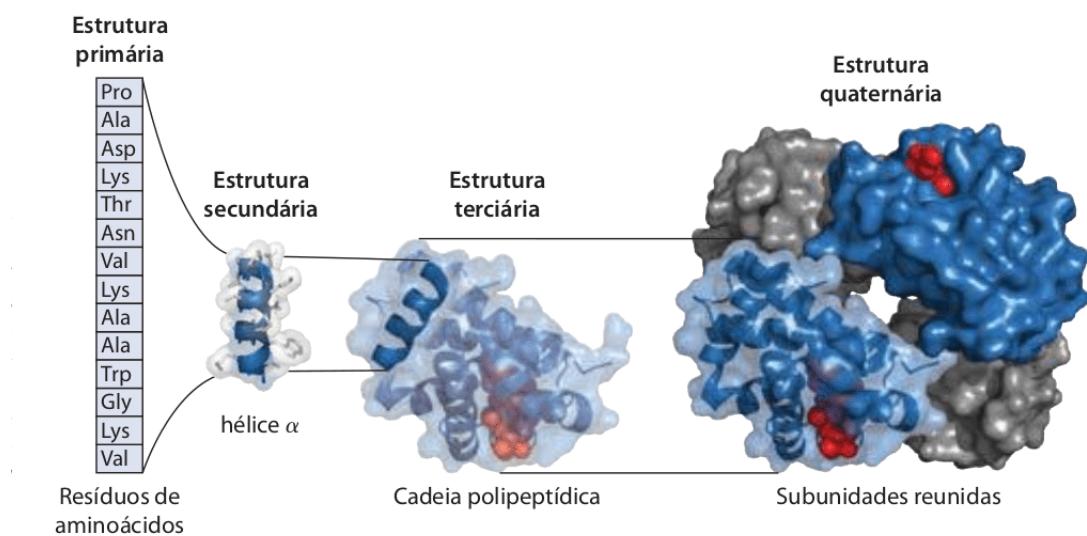


Figura 7 – Níveis de estrutura. Fonte: Extraída de (NELSON; COX, 2014)

- Estrutura Primária: cadeia linear que possui somente ligações peptídicas.

- Estrutura Secundária: α -hélices, fitas β e voltas são formadas por meio de um padrão regular de pontes de hidrogênio entre os grupos N—H e C=O dos aminoácidos que estão próximos uns dos outros na sequência linear do peptídeo ([BERG; TYMOCZKO; STRYER, 2014](#)).
- Estrutura Terciária: formas obtidas a partir das estruturas secundárias.
- Estrutura Quaternária: formas obtidas a partir das estruturas terciárias.

2.5 Algoritmos Evolutivos

Os Algoritmos Evolucionários(AEs) usam modelos computacionais baseados na evolução natural, segundo a teoria de Darwin, como o objetivo de resolver problemas.

Eles funcionam mantendo uma população de estruturas que evoluem de forma semelhante à evolução das espécies. A estas estruturas são aplicados os chamados operadores genéticos, como recombinação e mutação, entre outros. Cada indivíduo da população recebe uma avaliação que é uma quantificação da sua qualidade como solução do problema em abordado, e baseado nesta avaliação são aplicados operadores de forma a simular a sobrevivência do mais apto.([LINDEN, 2008](#))

Além da grande variedade de modelos computacionais propostos, todos eles têm em comum o conceito de simulação da evolução das espécies através de seleção, mutação e reprodução, processos estes que dependem do desempenho dos indivíduos desta espécie dentro do “ambiente”.([LINDEN, 2008](#))

Como exemplo desses algoritmos temos os Algoritmos Genéticos(AGs), que são consideradas técnicas heurísticas de otimização global([LINDEN, 2008](#)). Eles funcionam realizando uma busca não determinística abstraindo o conceito da Evolução.

O conceito é criar uma “população” inicial de estruturas, cada uma caracterizada por um conjunto de “genes”. As estruturas “pai” podem gerar “filhos” com uma mistura dos genes pai, permitindo que “mutações” ocorram no processo. As melhores espécies de uma população são selecionadas com base no princípio de Darwin, a sobrevivência do mais apto, e continuam para a próxima “geração”, enquanto as estruturas menos ajustadas são descartadas.([JENSEN, 2007](#))

3 Desenvolvimento

Este capítulo descreve o processo de estratégia, implementação e experimentos realizados com base na aplicação que foi desenvolvida utilizando a linguagem de programação Python na versão 3.6. Também especifica a máquina em que foi executado o programa e as bibliotecas de terceiros.

3.1 Sistema Operacional e Hardware

O Sistema Operacional (SO) utilizado foi o macOS High Sierra 10.13.6. O computador usado para execução da aplicação é um MacBook Air (13-inch, 2017) com processador 1,8 GHz Intel Core i5, 8 GB 1600 MHz DDR3 de memória RAM e 128 GB de espaço em disco SSD.

3.2 Bibliotecas de Terceiros

3.2.1 Psi4NumPy

É um módulo em python baseado na biblioteca Psi4 que possui implementação em C/C++.

O *Psi4NumPy* fornece uma estrutura de química quântica interativa para implementações de referência, prototipagem rápida, desenvolvimento e educação. Para fazer isso, as quantidades relevantes para a química quântica são calculadas com o pacote de estrutura eletrônica *Psi4* e, posteriormente, manipuladas usando o pacote *Numérico Python* (*NumPy*). Essa combinação fornece uma interface que é simples de usar e relativamente rápida de executar. (SMITH et al., 2018)

A instalação e utilização da biblioteca estão nos anexos A e B respectivamente.

3.2.2 SciPy

O ecossistema SciPy inclui ferramentas gerais e especializadas para gerenciamento e computação de dados, experimentação produtiva e computação de alto desempenho. (JONES et al., 2001–)

3.3 Estratégia para a modelagem

Foram utilizados arquivos contendo a estrutura 3D de 20 aminoácidos conhecidos, extraídos da base de dados PubChem.(KIM et al., 2018)

Validamos a sequência de entrada com o alfabeto de aminoácidos conhecidos e em seguida é feita a condensação dos mesmos, dando origem ao peptídeo resultante e eliminando a molécula de água (desidratação).

Por padrão, determinamos que a ligação peptídica entre os aminoácidos possui 1.32 Å, pois é a distância de ligação peptídica típica entre um Carbono(C) e um Nitrogênio(N).(SKERN, 2018)

A sequência de entrada é uma referência de como os aminoácidos serão condensados, pois a cadeia peptídica é formada por sequências de aminoácidos e a ordem deve ser respeitada. Por exemplo, se é digitado uma sequência “**AG**”, isso significa que é uma Alanina seguida por uma Glicina, invertendo isso é formado um peptídeo diferente.

Como existe uma posição no espaço 3D pré-definida para cada aminoácido, precisamos fazer o ajuste da posição dos aminoácidos subsequentes antes que a condensação ocorra. Para isso, utilizamos a **translação** da matriz de posições com base no vetor resultante entre dois átomos.

O vetor resultante entre os pontos a e b é dado por

$$\begin{aligned} a &= (x_a, y_a, z_a) \\ b &= (x_b, y_b, z_b) \\ v &= (x_a - x_b, y_a - y_b, z_a - z_b) \end{aligned} \tag{3.1}$$

Seja um ponto $P(x,y,z)$ sobre o qual será efetuada uma operação de translação e seja P' as coordenadas do ponto após a translação. Podemos definir a função T como

$$T(P) = T(x_p, y_p, z_p) = (x_p + d_x, y_p + d_y, z_p + d_z) \tag{3.2}$$

$$P' = P + T, \text{ onde } T = \begin{bmatrix} d_x \\ d_y \\ d_z \end{bmatrix} \tag{3.3}$$

Na figura 8, temos um exemplo de como fazemos o ajuste das posições para a condensação de dois aminoácidos. Com as posições 3D dos dois aminoácidos, calculamos o vetor resultante entre os átomos que vão ficar ligados pela ligação peptídica (figura 8a). E em seguida, realizamos a translação da matriz de posições do segundo aminoácido e a condensação entre eles (figura 8b), formando assim um dipeptídeo.

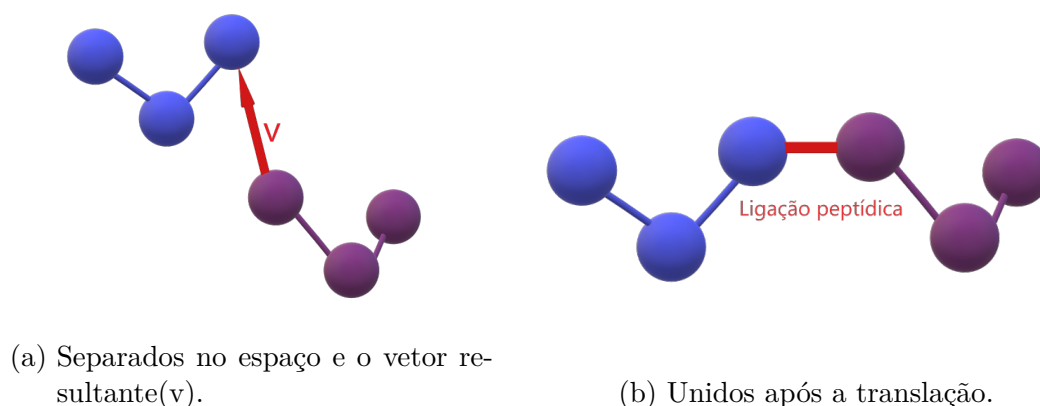


Figura 8 – Exemplo de dois aminoácidos formando um dipeptídeo

Outro exemplo é a condensação de três aminoácidos (figuras 9 e 10). Com as posições 3D pré-definidas temos os três aminoácidos em qualquer posição do espaço (figura 9a). Fazemos a condensação entre o primeiro e o segundo aminoácido seguindo o mesmo exemplo anterior, não considerando os aminoácidos seguintes para a translação (figura 9b).

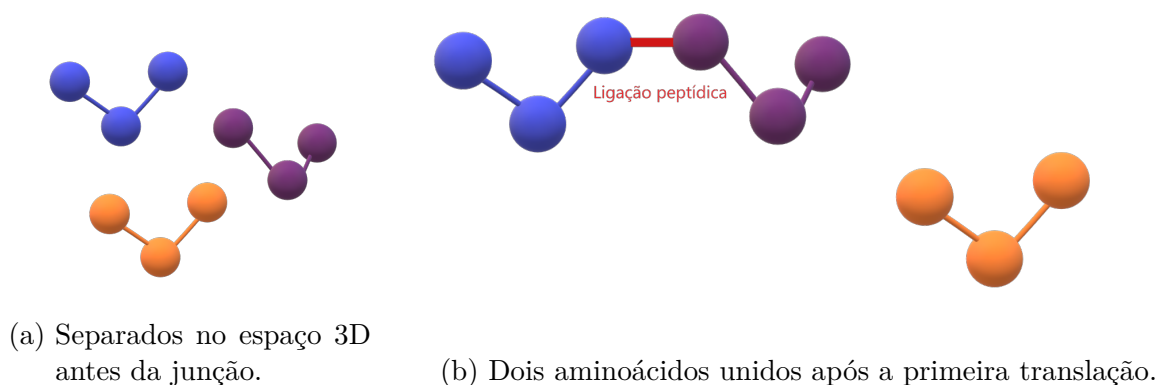


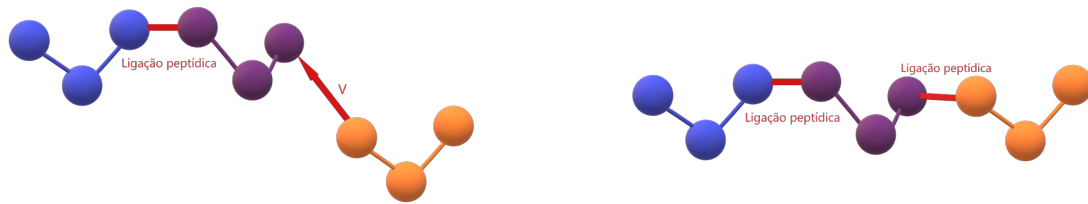
Figura 9 – Exemplo 1 de três aminoácidos formando um tripeptídeo

Em seguida, calculamos o vetor resultante entre os átomos que vão ficar ligados na segunda ligação peptídica (figura 10a). E finalmente, realizamos a translação da matriz de posições do terceiro aminoácido e a condensação com o dipeptídeo gerado anteriormente (figura 10b), formando assim um tripeptídeo.

Se a quantidade de aminoácidos for maior, continuamos o processo sucessivamente até que o peptídeo desejado seja formado.

Após a formação do peptídeo em sua conformação inicial, aplicamos o Algoritmo Genético(AG) com o objetivo de obtermos a conformação de menor energia.

Na mutação do Algoritmo Genético(AG), fazemos a rotação de parte da molécula em torno da ligação peptídica, ou seja, realizamos a torção. Se a molécula possui mais



- (a) Calcula o vetor unitário(v) para a segunda translação. (b) Três aminoácidos unidos após a segunda translação.

Figura 10 – Exemplo 2 de três aminoácidos formando um tripeptídeo

que uma ligação peptídica, rotacionamos toda a parte de um lado da ligação, que no caso do exemplo da figura 11 são as partes roxa e laranja. E em seguida, passamos para a próxima ligação peptídica e rotacionamos a parte seguinte da ligação, que é a parte laranja. Fazemos isso até que não haja mais ligações peptídicas. Como resultado disso, temos uma conformação diferente para o mesmo peptídeo resultando em uma energia diferente para ser avaliada.

Algoritmo Genético

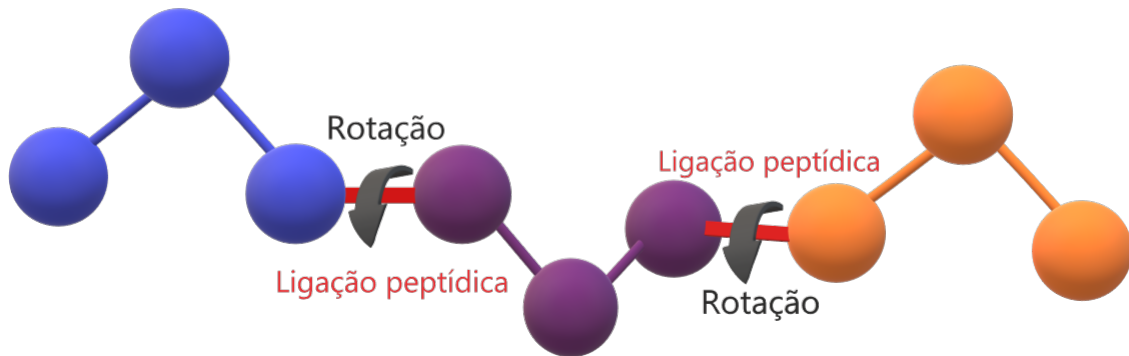


Figura 11 – Exemplo de rotações das ligações peptídicas utilizadas no Algoritmo Genético(AG).

3.4 Implementação

O programa recebe uma entrada de uma sequência de aminoácidos utilizando as letras de abreviação. Abaixo segue um exemplo de entrada do programa, nesse caso, a entrada **AG** significa uma Alanina(A) seguida de uma Glicina(G).

Listagem 3.1 – Exemplo de entrada

```
1 Enter an amino acid sequence: AG
```

A entrada é analisada verificando se cada caractere existe entre os 20 aminoácidos utilizados como base do programa. Se não tiver nenhum erro, buscamos as coordenadas de cada aminoácido e suas ligações fazendo a leitura dos arquivos com os dados dos aminoácidos extraídos da base de dados PubChem.

Após cada aminoácido ser instanciado, fazemos o processo de condensação seguindo a ideia explicada na estratégia para a modelagem. Fazemos a translação utilizando a seguinte implementação.

Listagem 3.2 – code.py

```
1 def resulting_vector(a, b):
2     ab = [a[0] - b[0], a[1] - b[1], a[2] - b[2]]
3     return ab
4
5 def resulting_vector_with_distance(a, b, distance):
6     v = resulting_vector(a, b)
7     v[0] = (a[0] - v[0]) + distance
8     v[1] = (a[1] - v[1]) + distance
9     v[2] = (a[2] - v[2]) + distance
10    return v
11
12 def translation(m, xyz):
13     for v in m:
14         v[0] = v[0] + xyz[0]
15         v[1] = v[1] + xyz[1]
16         v[2] = v[2] + xyz[2]
17    return m
```

Na condensação eliminamos água, e precisamos remover uma hidroxila(OH) de um aminoácido e um hidrogênio do outro, fazendo o ajuste das posições a casa ligação peptídica formada.

Listagem 3.3 – code.py

```
1 @staticmethod
2 def generate_peptide_molecule(amino_acid_sequence):
3     peptide = []
4
5     for a in amino_acid_sequence:
6         amino_acid = AminoAcid.generate_amino_acid_molecule(a)
7         if len(peptide) > 0:
8             last_amino_acid: AminoAcid = peptide[-1]
9             carbon = last_amino_acid.eliminate_hydroxyl_radical()
10            amino_acid.eliminate_hydrogen(carbon)
11            amino_acid = Molecule._update_position(last_amino_acid.
12                atoms[-1], amino_acid)
13            peptide.append(amino_acid)
```

```

13
14     return peptide
15
16 @staticmethod
17 def __update_position(atom_a: Atom, amino_acid: AminoAcid):
18     matrix = []
19     for atom in amino_acid.atoms:
20         matrix.append(atom.position)
21
22     v = resulting_vector_with_distance(atom_a.position, matrix[0],
23         PEPTIDE_BOND_LENGTH)
24
25     result = translation(matrix, v)
26
27     # update positions
28     for i in range(len(amino_acid.atoms)):
29         amino_acid.atoms[i].position = result[i]
30
31     return amino_acid

```

Com o peptídeo formado, podemos criar a população inicial e executar o Algoritmo Genético(AG) seguindo o fluxo da figura 12.

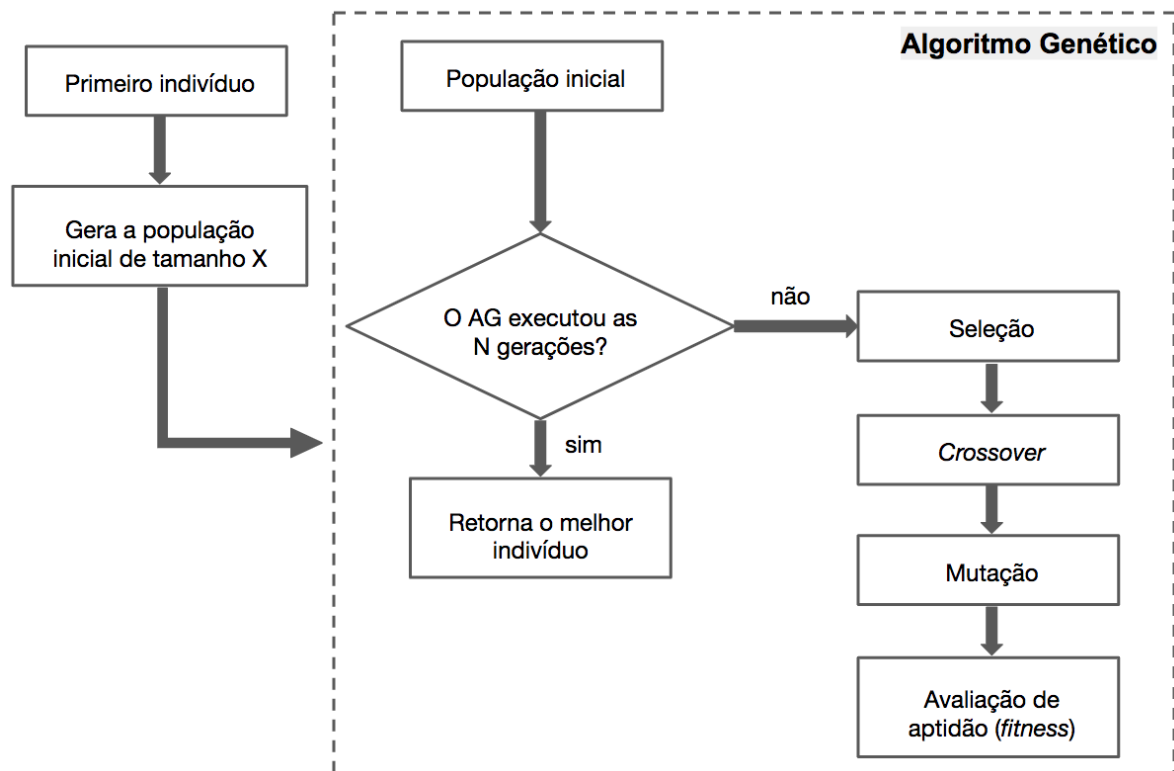


Figura 12 – Fluxograma do Algoritmo Genético(AG) implementado no programa.

A sele o   feita ordenando os indiv duos pela energia, e se a popula o estiver maior do que o tamanho estipulado, s o eliminados os indiv duos com maior energia.

O *fitness* é calculado baseado na energia eletrônica total do indivíduo, e buscamos sempre o de menor energia. Essa energia é calculada utilizando a biblioteca Psi4, com a configuração de cálculo base para otimização “**scf/cc-pvdz**”, que é o conjunto de funções base utilizada no método HF-SCF. Também é possível limitar a memória do computador que será utilizada nesse cálculo, que nesse caso foi de 500MB. Isso quer dizer que, se o cálculo da energia não convergir até os dados armazenados na memória chegarem ao limite máximo de 500MB, então a execução do cálculo encerra. Exemplo de utilização da biblioteca no anexo B.

A mutação é feita variando a torção em torno da ligação peptídica conforme a implementação abaixo.

Listagem 3.4 – mutation.py

```

1 def mutate(self):
2     if len(self.chromosome) == 1:
3         return
4
5     for i in range(1, len(self.chromosome)):
6         first_part: [AminoAcid] = self.chromosome[:i]
7         second_part: [AminoAcid] = self.chromosome[i:]
8
9         part_to_rotate = []
10        if len(second_part) == 1:
11            items = reduce(lambda x, y: x + y, second_part)
12            for atom in items.atoms:
13                part_to_rotate.append(atom.position)
14        else:
15            atoms = []
16            for x in second_part:
17                atoms.append(x.atoms)
18
19            items = reduce(lambda x, y: x + y, atoms)
20
21            for atom in items:
22                part_to_rotate.append(atom.position)
23
24        # calculates resulting vector
25        v = calculates_xyz_to_rotate(first_part[-1].atoms[-1].
26                                   position, part_to_rotate[0])
27
28        # move positions
29        new_part = rotation_euler(part_to_rotate, v)
30
31        # update values by reference
32        for j in range(len(second_part)):
33            for k in range(len(second_part[j].atoms)):

```

```
33         second_part[j].atoms[k].position = new_part[0]
34         del new_part[0]
35
36     # calculate fitness
37     self.fitness = self.__calculates_fitness()
38     print( Fitness:    + str(self.fitness))
```

O *crossover* é uniforme. Fazemos a seleção de dois indivíduos da população para gerar um filho onde os genes são a combinação das coordenadas de cada átomo dos peptídeos dos pais feitas de forma aleatória.

Listagem 3.5 – crossover.py

```
1 def __crossover(self):
2     offsprings = []
3     for i in range(0, POPULATION_SIZE-1, 2):
4         offspring = self.__uniform_crossover(self.population[i], self
5             .population[i+1])
6         offsprings.append(offspring)
7     self.population += offsprings
8
9 @staticmethod
10 def __uniform_crossover(parent_a: Individual, parent_b: Individual):
11     genes = []
12     for i in range(len(parent_a.chromosome)):
13         choice = random.randrange(2)
14         if choice == 0:
15             genes.append(parent_a.chromosome[i])
16         else:
17             genes.append(parent_b.chromosome[i])
18     return Individual(genes)
```

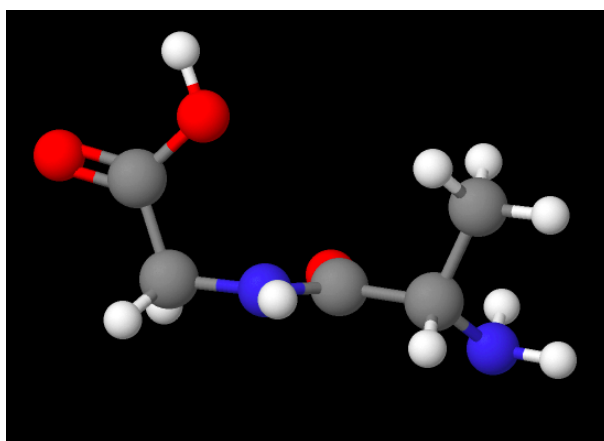
Após o AG rodar um número determinado de gerações, é selecionado o indivíduo com a menor energia para retornar como resultado final. Esse indivíduo contém a matriz de posições 3D utilizada para gerar a imagem da molécula a partir do arquivo de extensão “.mol”.

Como não encontrei nenhuma forma genérica para gerar o arquivo com as informações necessárias para gerar a imagem, utilizei o arquivo da base de dados PubChem e alterei manualmente a matriz de posições de cada átomo e as respectivas ligações entre eles. Após o arquivo ser atualizado, eu utilizei o site “<http://www.cheminfo.org>” para gerar as imagens a partir do arquivo “.mol”.

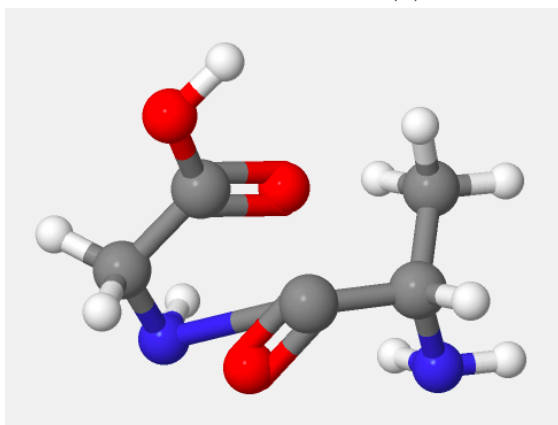
3.5 Resultados

Realizamos o cálculo da energia da molécula Ala-Gly a partir da estrutura 3D conhecida pela base de dados PubChem (Figura 13a), com auxílio da biblioteca Psi4 para comparar com o resultado obtido através do programa. A energia eletrônica total da estrutura conhecida possui o valor de **-528.7141214858434 kcal/mol** e a biblioteca Psi4 executou em **14.66 segundos**.

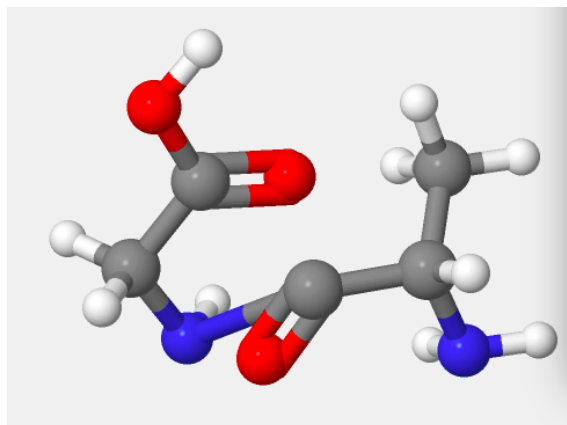
Ao executar o programa desenvolvido nesse trabalho e inserir a entrada “AG”, que é a forma abreviada de Ala-Gly, criou-se uma população inicial com cinco indivíduos que passaram pelo processo de mutação e *crossover* por cinco gerações. O melhor *fitness* para essa execução com um tempo de **16.67 minutos** foi o de **-528.323779795162 kcal/mol**, que é um valor bem próximo do obtido na execução anterior, e foram necessárias **13 interações** de HF-SCF para o cálculo da energia convergir. Podemos ver a conformação resultante na Figura 13b.



(a) Base de dados PubChem.



(b) Resultante do programa.



(c) Resultante da biblioteca Psi4.

Figura 13 – Projeção 3D da molécula Ala-Gly.

A biblioteca Psi4 gera um arquivo com vários dados, e também com a matriz 3D que sofre uma pequena alteração durante os cálculos de energia, essa alteração não

modifica significativamente os ângulos entre os átomos, e ao plotar a molécula temos a Figura 13c.

Há algumas diferenças geométricas entre as moléculas das figuras 13a e 13b, como a posição do grupos funcional Ácido Carboxílico(COOH), que pode mudar dependendo da torção que fazemos em torno de uma ligação simples.

As coordenadas que utilizamos para gerar as imagens das moléculas foram normalizadas seguindo a normalização de vetores, com o objetivo de obtermos o vetor unitário, que é o vetor cujo comprimento é 1 (equação 3.4).

$$u = \frac{v}{\|v\|}, \text{ onde } \|v\| = \sqrt{x^2 + y^2 + z^2} \quad (3.4)$$

Átomo	x	y	z
O ₁	0.7247	0.6477	-0.2347
N ₂	-0.7710	0.6289	0.0995
C ₃	-0.8656	-0.1387	0.4810
C ₄	-0.7193	-0.6468	-0.2534
C ₅	0.9919	0.1199	0.0408
H ₆	-0.4543	-0.2214	0.8628
H ₇	-0.3649	-0.9229	-0.1224
H ₈	-0.8738	-0.4774	-0.0917
H ₉	-0.5574	-0.4459	-0.7002
H ₁₀	-0.9041	0.3939	0.1654
H ₁₁	-0.6588	0.6496	-0.3792
O ₁₂	0.7206	-0.6339	-0.2807
O ₁₃	0.3509	-0.6546	-0.6695
N ₁₄	0.6357	0.4335	-0.6385
C ₁₅	0.8578	0.0211	-0.5135
C ₁₆	0.7347	-0.4590	-0.4993
H ₁₇	0.9714	0.0785	-0.2240
H ₁₈	0.7866	-0.0343	-0.6164
H ₁₉	0.3746	0.3447	-0.8606
H ₂₀	0.5402	-0.8168	-0.2024

Tabela 3 – Resultande do programa.

Átomo	x	y	z
O ₁	0.4493	0.6046	0.6576
N ₂	0.9843	0.1728	0.0335
C ₃	0.9779	-0.0890	-0.1891
C ₄	0.8087	-0.5876	0.0230
C ₅	0.8907	0.4517	0.0505
H ₆	0.8252	-0.0630	-0.5612
H ₇	0.5347	-0.8353	-0.1274
H ₈	0.8320	-0.5507	-0.0661
H ₉	0.7295	-0.5656	0.3844
H ₁₀	0.9981	0.0395	-0.0470
H ₁₁	0.9358	0.1466	0.3203
O ₁₂	-0.7986	-0.5838	0.1457
O ₁₃	-0.9964	0.0317	0.0774
N ₁₄	-0.1528	0.3164	-0.9362
C ₁₅	-0.8087	0.4563	-0.3711
C ₁₆	-0.9982	-0.0585	0.0076
H ₁₇	-0.6908	0.3884	-0.6098
H ₁₈	-0.6280	0.7775	-0.0322
H ₁₉	0.0159	-0.1314	-0.9911
H ₂₀	-0.7582	-0.6045	0.2440

Tabela 4 – Base PubChem.

Tabela 5 – Matriz das coordenadas da molécula Ala-Gly

$$\rho = \frac{cov(X,Y)}{\sqrt{var(X) \times var(Y)}} \quad (3.5)$$

Comparamos os coeficientes de correlação de Pearson (equação 3.5) para cada eixo do resultado do programa com a de referência e obtivemos os seguintes coeficientes:

- eixo x = -0.86

- eixo $y = 0.75$
- eixo $z = -0.01$

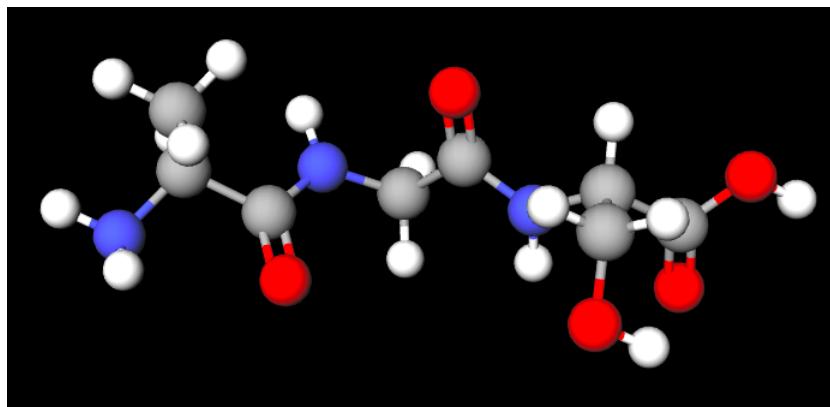
Para o dipeptídeo Ala-Gly, existe uma forte correlação para os eixos x e y . Já o eixo z não possui correlação.

Átomo	Átomo	Ligação
O_1	C_5	dupla
N_2	C_3	simples
N_2	H_{10}	simples
N_2	H_{11}	simples
C_3	C_4	simples
C_3	C_5	simples
C_3	H_6	simples
C_4	H_7	simples
C_4	H_8	simples
C_4	H_9	simples
O_{12}	C_{16}	simples
O_{12}	H_{20}	simples
O_{13}	C_{16}	dupla
N_{14}	C_{15}	simples
N_{14}	C_5	peptídica
N_{14}	H_{19}	simples
C_{15}	H_{17}	simples
C_{15}	H_{18}	simples
C_{16}	C_{15}	simples

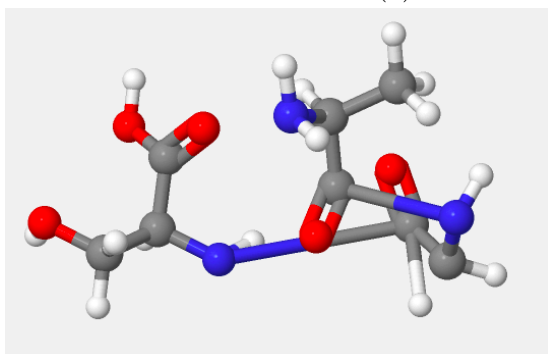
Tabela 6 – Ligações da molécula Ala-Gly

Realizamos também o teste do programa com o peptídeo Ala-Gly-Ser, composto por três aminoácidos que são Alanina, Glicina e Serina respectivamente. O resultado da execução do Psi4 para calcular a energia eletrônica total da molécula conhecida, extraída da PubChem(Figura 14a), foi de **-849.4371377028903680 kcal/mol** e o tempo de **47.63 segundos**.

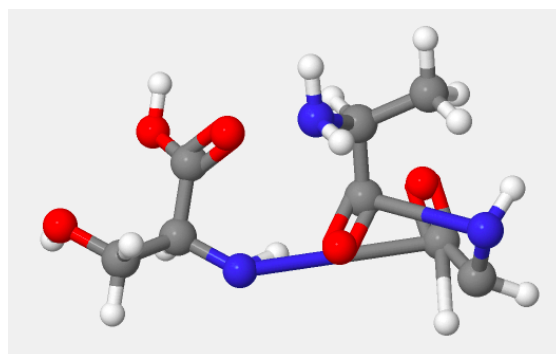
Ao executar o programa desenvolvido nesse trabalho e inserir a entrada “**AGS**”, que é a forma abreviada de Ala-Gly-Ser, criou-se uma população inicial com cinco indivíduos que passaram pelo processo de mutação e *crossover* por cinco gerações. O melhor *fitness* para essa execução foi o de **-848.7900243556691 kcal/mol**, que também é um valor bem próximo do obtido na execução anterior, com um tempo de **55.80 minutos**. Podemos ver a conformação resultante na Figura 14b. E foram necessárias **23 interações** de HF-SCF para o cálculo da energia convergir.



(a) Base de dados PubChem.



(b) Resultante do programa.



(c) Resultante da biblioteca Psi4.

Figura 14 – Projeção 3D da molécula Ala-Gly-Ser.

A figura 14a, que utilizamos como base, possui uma cadeia mais linear do que a figura 14b que foi resultante e alguns comprimentos de ligação sofreram deformação por causa da mutação durante a execução do algoritmo genético.

As coordenadas que utilizamos para gerar as imagens das moléculas também foram normalizadas seguindo a normalização de vetores (equação 3.4). E assim obtemos os dados da tabela 9.

Comparamos os coeficientes de correlação de Pearson (equação 3.5) para cada eixo do resultado do programa com a de referência e obtivemos os seguintes coeficientes:

- eixo x = -0.01
- eixo y = -0.03
- eixo z = -0.04

Para o tripeptídeo Ala-Gly-Ser, não existe correlação para nenhum dos eixos.

Por último, realizamos também o teste do programa com o peptídeo Ala-Gly-Ser-Glu, composto por três aminoácidos que são Alanina, Glicina, Serina e Ácido Glutâmico respectivamente. O resultado da execução do Psi4 para calcular a energia eletrônica total

Átomo	x	y	z
O ₁	0.7247	0.6477	-0.2347
N ₂	-0.7710	0.6289	0.0995
C ₃	-0.8656	-0.1387	0.4810
C ₄	-0.7193	-0.6468	-0.2534
C ₅	0.9919	0.1199	0.0408
H ₆	-0.4543	-0.2214	0.8628
H ₇	-0.3649	-0.9229	-0.1224
H ₈	-0.8738	-0.4774	-0.0917
H ₉	-0.5574	-0.4459	-0.7002
H ₁₀	-0.9041	0.3939	0.1654
H ₁₁	-0.6588	0.6496	-0.3792
O ₁₂	0.5694	-0.8191	-0.0689
N ₁₃	0.2064	-0.2240	-0.9524
C ₁₄	0.6327	-0.2693	-0.7259
C ₁₅	0.7984	-0.5075	-0.3237
H ₁₆	0.7300	0.0285	-0.6828
H ₁₇	0.5922	-0.3975	-0.7008
H ₁₈	0.0132	-0.5066	-0.8620
O ₁₉	0.1805	0.7676	0.6149
O ₂₀	0.0708	0.4020	0.9128
O ₂₁	-0.1466	0.5860	0.7969
N ₂₂	0.6373	0.6527	0.4094
C ₂₃	0.4120	0.6531	0.6352
C ₂₄	0.2944	0.8186	0.4930
C ₂₅	0.1187	0.5566	0.8222
H ₂₆	0.5058	0.5099	0.6958
H ₂₇	0.4212	0.8191	0.3892
H ₂₈	0.1524	0.9149	0.3737
H ₂₉	0.7675	0.4094	0.4932
H ₃₀	0.2746	0.6881	0.6715
H ₃₁	-0.1002	0.3041	0.9473

Tabela 7 – Resultande do programa.

Átomo	x	y	z
O ₁	0.8587	-0.4275	0.2825
N ₂	0.9826	-0.1563	0.0999
C ₃	0.9991	0.0332	-0.0240
C ₄	0.9343	-0.0074	-0.3562
C ₅	0.9870	-0.0342	0.1567
H ₆	0.9690	0.2447	0.0336
H ₇	0.8507	0.1643	-0.4991
H ₈	0.9324	0.0215	-0.3605
H ₉	0.8736	-0.2318	-0.4278
H ₁₀	0.9936	-0.1094	0.0276
H ₁₁	0.9393	-0.3396	0.0472
O ₁₂	0.0015	0.6966	-0.7174
N ₁₃	0.8680	0.4467	0.2164
C ₁₄	0.4363	0.6644	0.6067
C ₁₅	-0.2497	0.9597	-0.1284
H ₁₆	0.3156	0.1842	0.9308
H ₁₇	0.2135	0.8043	0.5544
H ₁₈	0.7876	0.6132	0.0592
O ₁₉	-0.7305	-0.6825	-0.0197
O ₂₀	-0.9840	0.1486	-0.0979
O ₂₁	-0.9379	-0.0126	0.3464
N ₂₂	-0.9512	0.2222	0.2137
C ₂₃	-0.9769	0.0233	-0.2121
C ₂₄	-0.8155	-0.4429	-0.3724
C ₂₅	-0.9955	0.0629	0.0698
H ₂₆	-0.8487	0.2685	-0.4555
H ₂₇	-0.5672	-0.5373	-0.6241
H ₂₈	-0.8166	-0.3823	-0.4322
H ₂₉	-0.7573	0.0397	0.6518
H ₃₀	-0.8212	-0.5633	0.0901
H ₃₁	-0.9894	0.1443	0.0079

Tabela 8 – Base PubChem.

Tabela 9 – Matriz normalizada das coordenadas da molécula Ala-Gly-Ser

da molécula conhecida, extraída da PubChem(Figura 15a), foi de **-1321.9687202425043 kcal/mol** e o tempo de **2.05 minutos**.

Executamos o programa com a entrada “**AGSE**”, que é a forma abreviada de Ala-Gly-Ser-Glu, criou-se uma população inicial com cinco indivíduos que passaram pelo processo de mutação e *crossover* por cinco gerações. O melhor *fitness* para essa execução foi o de **-1320.0956240984865 kcal/mol**, que também é um valor bem próximo do obtido na execução anterior, com um tempo de **7.11 horas**. Foram necessárias **20 interações** de HF-SCF para o cálculo da energia convergir.

Átomo	Átomo	Ligação
O_1	C_5	dupla
N_2	C_3	simples
N_2	H_{10}	simples
N_2	H_{11}	simples
C_3	C_4	simples
C_3	C_5	simples
C_3	H_6	simples
C_4	H_7	simples
C_4	H_8	simples
C_4	H_9	simples
O_{12}	C_{15}	dupla
N_{13}	C_{14}	simples
N_{13}	C_5	peptídica
N_{13}	H_{18}	simples
C_{14}	C_{15}	simples
C_{14}	C_{16}	simples
C_{14}	H_{17}	simples
O_{19}	C_{24}	simples
O_{19}	H_{30}	simples
O_{20}	C_{25}	simples
O_{20}	H_{31}	simples
O_{21}	C_{25}	dupla
O_{21}	C_{23}	simples
N_{22}	C_{15}	peptídica
N_{22}	H_{29}	simples
C_{23}	C_{24}	simples
C_{23}	C_{25}	simples
C_{23}	H_{26}	simples
C_{24}	H_{27}	simples
C_{24}	H_{28}	simples

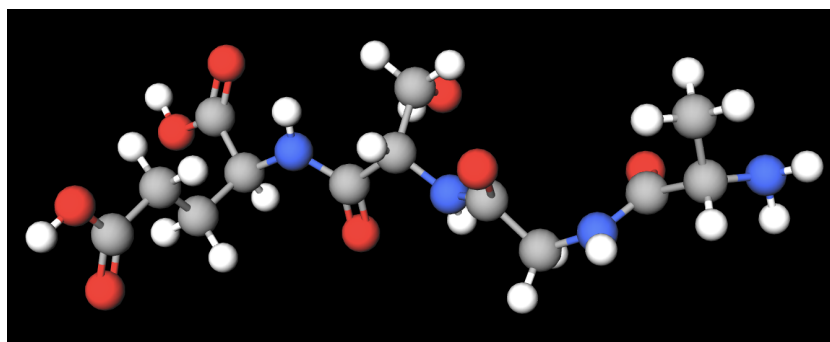
Tabela 10 – Ligações da molécula Ala-Gly-Ser

Podemos observar uma forte deformação da molécula no resultado do programa, figura 15b. Isso devido ao processo de *crossover* que é uma das etapas do Algoritmo Genético utilizada para gerar novos indivíduos. Apesar disso, a energia eletrônica total é bem próxima comparada com a energia da molécula de referência.

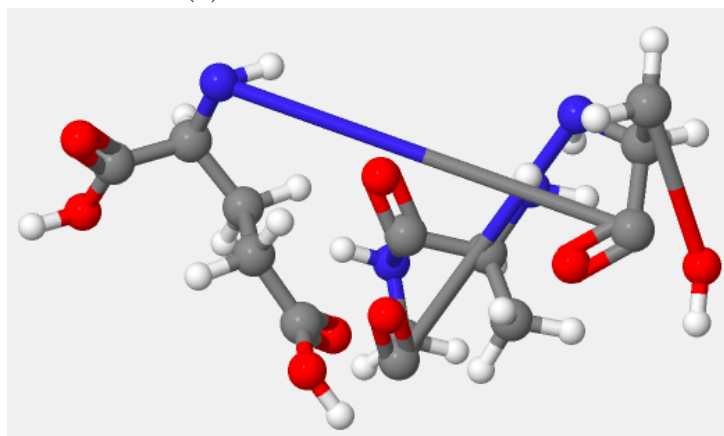
As coordenadas que utilizamos para gerar as imagens das moléculas também foram normalizadas seguindo a normalização de vetores (equação 3.4). E assim obtemos os dados da tabela 13.

Comparamos os coeficientes de correlação de Pearson para cada eixo do resultado do programa com a de referência e obtivemos os seguintes coeficientes:

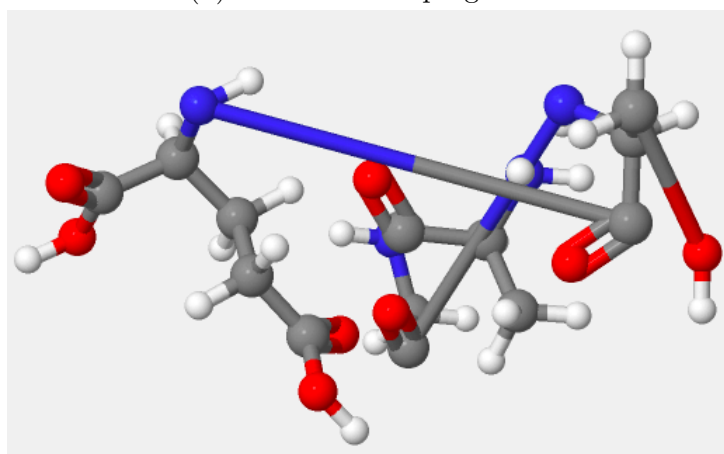
- eixo x = 0.57



(a) Base de dados PubChem.



(b) Resultante do programa.



(c) Resultante da biblioteca Psi4.

Figura 15 – Projeção 3D da molécula Ala-Gly-Ser-Glu.

- eixo $y = -0.58$
- eixo $z = -0.59$

Para o tetrapeptídeo Ala-Gly-Ser-Glu, existe uma forte correlação moderada para os eixos x , y e z .

Todos os logs de saída da execução do programa estão no anexo C.

Átomo	x	y	z
O ₁	0.7247	0.6477	-0.2347
N ₂	-0.7710	0.6289	0.0995
C ₃	-0.8656	-0.1387	0.4810
C ₄	-0.7193	-0.6468	-0.2534
C ₅	0.9919	0.1199	0.0408
H ₆	-0.4543	-0.2214	0.8628
H ₇	-0.3649	-0.9229	-0.1224
H ₈	-0.8738	-0.4774	-0.0917
H ₉	-0.5574	-0.4459	-0.7002
H ₁₀	-0.9041	0.3939	0.1654
H ₁₁	-0.6588	0.6496	-0.3792
O ₁₂	0.5689	-0.8194	-0.0685
N ₁₃	0.4463	-0.2104	0.8697
C ₁₄	0.2760	-0.6384	0.7184
C ₁₅	0.3026	-0.9044	0.3006
H ₁₆	-0.0338	-0.6250	0.7798
H ₁₇	0.3947	-0.6482	0.6510
H ₁₈	0.7294	-0.1377	0.6699
O ₁₉	-0.9565	0.0286	-0.2900
O ₂₀	-0.8251	-0.0137	-0.5647
N ₂₁	-0.6352	0.7510	-0.1798
C ₂₂	-0.8030	0.5111	-0.3063
C ₂₃	-0.6655	0.5452	-0.5096
C ₂₄	-0.9036	0.1748	-0.3910
H ₂₅	-0.8600	0.4919	-0.1349
H ₂₆	-0.5936	0.6743	-0.4390
H ₂₇	-0.5340	0.5162	-0.6695
H ₂₈	-0.7072	0.7005	0.0947
H ₂₉	-0.9370	-0.1698	-0.3052
O ₃₀	0.9982	-0.0072	-0.0586
O ₃₁	0.5343	-0.6138	-0.5811
O ₃₂	0.9291	0.1912	-0.3163
O ₃₃	0.7176	-0.6963	0.0118
N ₃₄	0.8338	0.5020	-0.2295
C ₃₅	0.9957	0.0838	-0.0387
C ₃₆	0.9549	0.2866	-0.0764
C ₃₇	0.9134	-0.1266	-0.3867
C ₃₈	0.9726	0.1495	-0.1777
C ₃₉	0.7711	-0.5266	-0.3578
H ₄₀	0.9610	0.2482	0.1210
H ₄₁	0.9862	-0.0959	0.1343
H ₄₂	0.9353	0.3369	0.1076
H ₄₃	0.7978	0.0655	-0.5992
H ₄₄	0.8923	-0.1889	-0.4098
H ₄₅	0.7500	0.6433	-0.1533
H ₄₆	0.9922	-0.0502	-0.1140
H ₄₇	0.3592	-0.7825	-0.5084

Tabela 11 – Resultante do programa.

Átomo	x	y	z
O ₁	-0.9776	-0.1810	-0.1068
N ₂	-0.9996	-0.0182	-0.0199
C ₃	-0.9935	0.1054	0.0417
C ₄	-0.9603	0.1017	0.2594
C ₅	-0.9974	0.0379	-0.0605
H ₆	-0.9702	0.2418	-0.0084
H ₇	-0.9466	-0.0437	0.3193
H ₈	-0.9498	0.1334	0.2828
H ₉	-0.9179	0.2117	0.3355
H ₁₀	-0.9993	0.0198	0.0290
H ₁₁	-0.9894	-0.0075	-0.1444
O ₁₂	-0.8164	0.4578	0.3518
N ₁₃	-0.9579	0.2584	-0.1250
C ₁₄	-0.9032	0.2784	-0.3263
C ₁₅	-0.9459	0.3213	-0.0431
H ₁₆	-0.8713	0.0437	-0.4887
H ₁₇	-0.7580	0.4817	-0.4395
H ₁₈	-0.9088	0.4120	-0.0643
O ₁₉	-0.2349	-0.9678	0.0900
O ₂₀	0.4613	-0.2540	-0.8500
N ₂₁	-0.9192	-0.0562	-0.3896
C ₂₂	-0.2389	-0.8186	0.5221
C ₂₃	-0.1987	-0.8186	0.5388
C ₂₄	0.7719	-0.4464	-0.4524
H ₂₅	0.0843	0.3793	0.9214
H ₂₆	0.1596	-0.7324	0.6618
H ₂₇	-0.4856	-0.5785	0.6552
H ₂₈	-0.6488	-0.2338	-0.7240
H ₂₉	0.0466	-0.9963	-0.0717
O ₃₀	0.9176	-0.3929	-0.0589
O ₃₁	0.8576	0.4698	0.2089
O ₃₂	0.8193	-0.5014	0.2777
O ₃₃	0.8787	0.4564	-0.1395
N ₃₄	0.9378	-0.3454	0.0345
C ₃₅	0.9807	0.0803	-0.1779
C ₃₆	0.9525	-0.2659	-0.1480
C ₃₇	0.9655	0.2384	0.1038
C ₃₈	0.9079	-0.4151	0.0568
C ₃₉	0.9069	0.4195	0.0379
H ₄₀	0.9034	0.2445	-0.3522
H ₄₁	0.9679	0.0311	-0.2491
H ₄₂	0.8470	-0.3671	-0.3844
H ₄₃	0.9062	0.3338	0.2592
H ₄₄	0.9730	0.0992	0.2082
H ₄₅	0.8515	-0.3124	0.4210
H ₄₆	0.8936	-0.4480	0.0257
H ₄₇	0.8294	0.5342	0.1631

Tabela 12 – Base PubChem.

Tabela 13 – Matriz normalizada das coordenadas da molécula Ala-Gly-Ser-Glu

Átomo	Átomo	Ligação
O_1	C_5	dupla
N_2	C_3	simples
N_2	H_{10}	simples
N_2	H_{11}	simples
C_3	C_4	simples
C_3	C_5	simples
C_3	H_6	simples
C_4	H_7	simples
C_4	H_8	simples
C_4	H_9	simples
O_{12}	C_{15}	dupla
N_{13}	C_{14}	simples
N_{13}	C_5	peptídica
N_{13}	H_{18}	simples
C_{14}	C_{15}	simples
C_{14}	C_{16}	simples
C_{14}	H_{17}	simples
O_{19}	C_{23}	simples
O_{19}	H_{29}	simples
O_{20}	C_{24}	dupla
N_{21}	C_{22}	simples
N_{21}	C_{29}	peptídica
N_{21}	H_{28}	simples
C_{22}	C_{23}	simples
C_{22}	C_{24}	simples
C_{22}	H_{25}	simples
C_{23}	H_{26}	simples
C_{23}	H_{27}	simples
O_{30}	C_{38}	simples
O_{30}	H_{46}	simples
O_{31}	C_{39}	simples
O_{31}	H_{47}	simples
O_{32}	C_{38}	dupla
O_{33}	C_{39}	dupla
N_{34}	C_{36}	simples
N_{34}	C_{24}	peptídica
N_{34}	H_{45}	simples
C_{35}	C_{36}	simples
C_{35}	C_{37}	simples
C_{35}	H_{40}	simples
C_{35}	H_{41}	simples
C_{36}	C_{38}	simples
C_{36}	H_{42}	simples
C_{37}	C_{39}	simples
C_{37}	H_{43}	simples
C_{37}	H_{44}	simples

Tabela 14 – Ligações da molécula Ala-Gly-Ser-Glu

Nos quatro experimentos com o programa, podemos observar que a conformação da molécula resultante não é exatamente igual à molécula de referência conhecida. Mesmo assim a energia eletrônica total está bem próxima, pois a biblioteca Psi4 recebe como parâmetro de entrada as coordenadas de cada átomo, e a energia de uma interação pode compensar a outra fazendo com que a energia total do sistema tenha resultados parecidos mesmo com conformações diferentes.

O tempo de execução do programa aumentou significativamente da entrada “AG” para a entrada “AGS”, e ainda mais para a entrada “AGSE”. Isso já era esperado, pois o problema pertence a NP-Completo e possui tempo exponencial, figura 16.

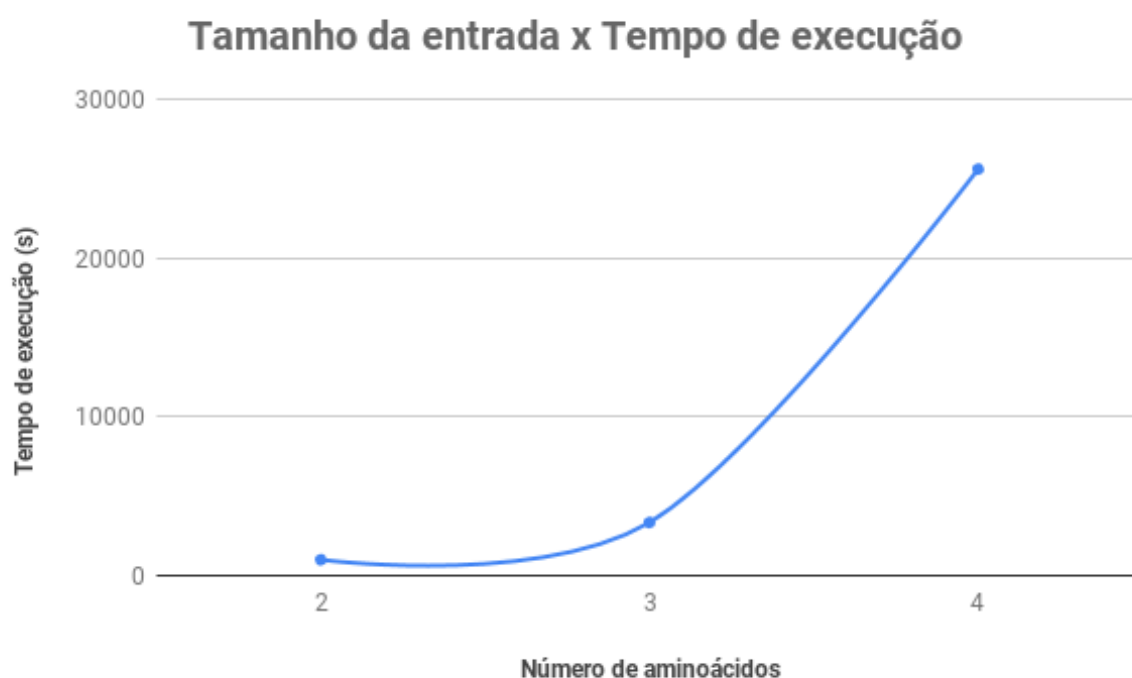


Figura 16 – Gráfico que mostra o tempo de execução para cada entrada do programa.

4 Conclusão

Ainda há muito a ser explorado nesta área de predição de estruturas tridimensionais(3D) de peptídeos e proteínas, principalmente pela complexidade dos cálculos envolvidos. Com o avanço dos Algoritmos Evolutivos(AEs) e suas diversas vertentes, podemos descobrir maneiras diferentes de tratar esse problema.

Utilizando os aminoácidos existentes, que são a base de qualquer peptídeo ou proteína no programa, temos uma molécula pré-definida que facilita muito para que não seja necessário fazer grandes modificações na matriz 3D. Podemos observar isso nos resultados obtidos com dois e três aminoácidos, pois o cálculo da energia eletrônica total ficou bem próximo ao de uma estrutura conhecida.

Para a continuação deste trabalho, é necessário ajustar o algoritmo de *crossover* para evitar a deformação da molécula e manter as distâncias de ligação.

Seria interessante, utilizar uma base de dados com as estruturas 3D maior para utilizar a maior cadeia conhecida do peptídeo, pois acredito que não seria necessário tantas interações e grandes modificações nos ângulos, pois já estariam bem próximos à sua forma estável e o resultado iria a convergir mais rápido. E também testar outros conjuntos de funções base no método HF-SCF para analisar o desempenho do tempo de processamento e nos resultados dos cálculos da energia.

Então, esse trabalho proporcionou uma visão do quão complexo é o problema da Modelagem Molecular e que existem muitos fatores que influenciam no resultado. E por fazer parte da classe dos problemas NP-Completo, torna o avanço nessa área bem mais lento porque o tempo computacional é um fator limitante e não possuímos computadores quânticos para resolvermos esse problema em tempo polinomial.

Referências

- AMINOÁCIDOS. In: WIKIPÉDIA: a enciclopédia livre. Wikimedia, 2017. Disponível em: <https://pt.wikipedia.org/wiki/Amino%C3%A1cido#cite_note-ucalgary.27.0-1>. Citado 3 vezes nas páginas 8, 9 e 20.
- AMUSIA, M.; MSEZANE, A. Z.; SHAGINYAN, V. The hartree fock method within density functional theory. 05 2003. Citado na página 16.
- BERG, J. M.; TYMOCZKO, J. L.; STRYER, L. *Bioquímica*. 7. ed. Rio de Janeiro: Guanabara Koogan, 2014. ISBN 978-8-5277-2387-9. Citado na página 22.
- BERGER, B.; LEIGHTON, T. Protein folding in the hydrophobic-hydrophilic (hp) model is np-complete. *Journal of Computational Biology*, v. 5, p. 27–40, 1998. Citado na página 14.
- CARVALHO, I. et al. Introdução a modelagem molecular de fármacos no curso experimental de química farmacêutica. *Química Nova*, scielo, v. 26, p. 428 – 438, 05 2003. ISSN 0100-4042. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-40422003000300023&nrm=iso>. Citado na página 16.
- CRESCENZI, P. et al. On the complexity of protein folding. *Journal of Computational Biology*, v. 5, n. 3, p. 423–465, 1998. Disponível em: <<https://doi.org/10.1089/cmb.1998.5.423>>. Citado na página 15.
- GARCIA CLEVERSON FERNANDO; FERREIRA LUCAS, E. M. B. I. *Química orgânica: Estrutura e propriedades*. 6. ed. Porto Alegre: Bookman Editora, 2015. 78 p. ISBN 8582602448, 9788582602447. Citado na página 19.
- GAUSSIAN Basis Sets. In: EXPANDING the limits of computational chemistry. Gaussian, 2017. Disponível em: <<https://gaussian.com/basissets>>. Citado 2 vezes nas páginas 9 e 17.
- ISOMERIA geométrica. In: WIKIPÉDIA: a enciclopédia livre. Wikimedia, 2018. Disponível em: <https://pt.wikipedia.org/wiki/Isomeria_geom%C3%A9trica>. Citado 2 vezes nas páginas 8 e 19.
- JENSEN, F. *Introduction to computational chemistry*. 2. ed. West Sussex PO19 8SQ, England: WILEY, 2007. ISBN 13 978-0-470-01186-7. Citado 2 vezes nas páginas 14 e 22.
- JONES, E. et al. *SciPy: Open source scientific tools for Python*. 2001–. Acessado em: 2019. Disponível em: <<http://www.scipy.org/>>. Citado na página 23.
- KIM, S. et al. PubChem 2019 update: improved access to chemical data. *Nucleic Acids Research*, v. 47, n. D1, p. D1102–D1109, 10 2018. ISSN 0305-1048. Disponível em: <<https://doi.org/10.1093/nar/gky1033>>. Citado na página 24.
- LEVINE, I. N. *Quantum Chemistry*. Englewood Cliffs, New Jersey: Prentice Hall: Springer, Dordrecht, 1991. ISBN 0-205-12770-3. Citado na página 17.

- LINDEN, R. *Algoritmos Genéticos: Teoria e Implementação*. 2. ed. Rio de Janeiro: BRASFORT, 2008. ISBN 978-85-7452-373-6. Citado na página 22.
- MAUPETIT, J.; DERREUMAUX, P.; TUFFERY, P. Pep-fold: an online resource for de novo peptide structure prediction. *Nucleic Acids Research*, v. 37, p. W498–W503, 2009. Disponível em: <<http://dx.doi.org/10.1093/nar/gks323>>. Citado na página 15.
- MÉTODOS ab initio. In: WIKIPÉDIA: a enciclopédia livre. Wikimedia, 2017. Disponível em: <https://pt.wikipedia.org/wiki/M%C3%A9todos_ab_initio>. Citado na página 17.
- MODELAGEM molecular. In: WIKIPÉDIA: a enciclopédia livre. Wikimedia, 2016. Disponível em: <https://pt.wikipedia.org/wiki/Modelagem_molecular>. Citado na página 16.
- NELSON, D. L.; COX, M. M. *Princípios de Bioquímica de Lehninger*. 6. ed. Porto Alegre: artmed, 2014. ISBN 978-85-8271-073-9. Citado 5 vezes nas páginas 8, 14, 19, 20 e 21.
- RAMACHANDRAN, K. I.; DEEPA, G.; NAMBOORI, K. *Computational Chemistry and Molecular Modeling: Principles and Applications*. Coimbatore, India: Springer-Verlag Berlin Heidelberg, 2008. ISBN 978-3-540-77302-3. Citado na página 17.
- RIGDEN, D. J. *From Protein Structure to Function with Bioinformatics*. 1. ed. United Kingdom: Springer, Dordrecht, 2017. ISBN 978-1-4020-9057-8. Citado na página 17.
- SKERN, T. *Exploring Protein Structure: Principles and Practice*. 1. ed. [S.l.]: Springer International Publishing, 2018. ISBN 978-3-319-76858-8. Citado na página 24.
- SMITH, D. G. A. et al. Psi4numpy: An interactive quantum chemistry programming environment for reference implementations and rapid development. *Journal of Chemical Theory and Computation*, v. 14, n. 7, p. 3504–3511, 2018. PMID: 29771539. Disponível em: <<https://doi.org/10.1021/acs.jctc.8b00286>>. Citado 2 vezes nas páginas 16 e 23.
- SOLOMONS, T. W. G.; FRYHLE, C. B. *Química Orgânica*. 7. ed. Rio de Janeiro: LTC, 2000. v. 1. ISBN 8521612826. Citado 2 vezes nas páginas 8 e 18.
- SUN, S. A genetic algorithm that seeks native states of peptides and proteins. *Biophysical journal*, Elsevier, v. 69, n. 2, p. 340–355, 1995. Citado na página 15.
- THÉVENET, P. et al. Pep-fold: an updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides. *Nucleic Acids Research*, v. 40, n. W1, p. W288–W293, 2012. Disponível em: <<http://dx.doi.org/10.1093/nar/gks419>>. Citado na página 15.

Anexos

ANEXO A – Instalação das bibliotecas necessárias para a utilização do programa

A.1 Psi4NumPy

Para instalar, basta seguir os seguintes passos:

Listagem A.1 – Terminal

```
1 # Have Psi4conda installer (http://psicode.org/downloads.html)
2 $ bash psi4conda-{various}.sh
3 # Check psi4 command in path; adjust path if needed
4 # **IF** using DFT tutorials (or a few newer specialized integrals),
   after above, create a separate environment within for newer psi4:
5 $ conda create -n p4env psi4 -c psi4/label/dev
6 $ source activate p4env
```

A.2 SciPy

Para instalar, basta digitar o seguinte comando no terminal:

Listagem A.2 – Terminal

```
1 $ python -m pip install numpy scipy matplotlib
```

ANEXO B – Utilização da biblioteca Psi4

Listagem B.1 – code.py

```
1 import psi4
2
3 total_energy = 0
4 geometry = """
5 C    0.000000   -0.667578   -2.124659
6 C    0.000000    0.667578   -2.124659
7 H    0.923621   -1.232253   -2.126185
8 H   -0.923621   -1.232253   -2.126185
9 H   -0.923621    1.232253   -2.126185
10 H    0.923621    1.232253   -2.126185
11 """
12 try:
13     psi4.core.set_output_file("output.dat", False)
14     psi4.set_memory("500 MB")
15
16     psi4.geometry(geometry)
17
18     psi4.energy("scf/cc-pvdz")
19     total_energy = psi4.core.get_variable("SCF TOTAL ENERGY")
20 except RuntimeError:
21     # Iterations did not converge.
22     pass
```

ANEXO C – Saídas do programa implementado nesse trabalho

C.1 Ala-Gly

Listagem C.1 – Log

```
1 Enter an amino acid sequence: AG
2
3 Generates Population
4
5 Fitness: -527.7803265267446
6 Fitness: -527.6776932763687
7 Fitness: -527.0522166171982
8 Fitness: -528.2659376789261
9 Fitness: -528.2807625975083
10
11 GENERATION 0
12 Fitness: -527.8325001357168
13 Fitness: -528.2046184793761
14 Fitness: -528.0332323598259
15 Fitness: -528.2720018154023
16 Fitness: -528.2651706839677
17 Fitness: -528.1628077206278
18 Fitness: -527.7780622610587
19
20 GENERATION 1
21 Fitness: -528.3422254907148
22 Fitness: -528.0304790449172
23 Fitness: -528.0620327079089
24 Fitness: -528.3400780939115
25 Fitness: -528.2396622020563
26 Fitness: -527.9126030997212
27 Fitness: -528.2111578094529
28
29 GENERATION 2
30 Fitness: -528.0686838504043
31 Fitness: -528.1264525525735
32 Fitness: -528.2706526403547
33 Fitness: -528.253510738619
34 Fitness: 0
35 Fitness: -527.3271472997753
36 Fitness: -528.2103042030304
```



```
37
38 GENERATION 3
39 Fitness: -528.0649666612246
40 Fitness: -528.2709404573242
41 Fitness: -528.268292892008
42 Fitness: -528.1680579018376
43 Fitness: -527.6182987927125
44 Fitness: -528.321031909251
45 Fitness: -527.9977306496576
46
47 GENERATION 4
48 Fitness: -528.2085618711468
49 Fitness: -528.323779795162
50 Fitness: -528.2327206560506
51 Fitness: -527.7966275124737
52 Fitness: -528.2936018701213
53 Fitness: -528.0277999845953
54 Fitness: -528.1257268005738
55
56
57 The best fitness: -528.323779795162
58
59 The best geometry:
60
61 O 1.2492 1.1165 -0.4047
62 N -1.4105 1.1507 0.1821
63 C -0.7085 -0.1136 0.3937
64 C -1.3345 -1.2000 -0.4702
65 C 0.7470 0.0903 0.0308
66 H -0.7666 -0.3737 1.4558
67 H -0.8580 -2.1695 -0.2878
68 H -2.4023 -1.3127 -0.2521
69 H -1.2248 -0.9797 -1.5384
70 H -2.3916 1.0420 0.4376
71 H -1.4071 1.3875 -0.8099
72 O 2.5516 -2.2446 -0.9942
73 O 0.5915 -1.1035 -1.1285
74 N 1.7707 1.2076 -1.7786
75 C 2.6493 0.0652 -1.5860
76 C 1.8114 -1.1318 -1.2310
77 H 3.3505 0.2709 -0.7726
78 H 3.2005 -0.1396 -2.5078
79 H 1.0947 1.0073 -2.5146
80 H 1.9913 -3.0108 -0.7463
81 THE END
82
83 Process finished with exit code 0
```

Listagem C.2 – output.dat

```

1  Memory set to 476.837 MiB by Python driver.
2
3  *** tstart() called on Tatianes-MacBook-Air.local
4  *** at Thu Jun 20 14:03:04 2019
5
6  => Loading Basis Set <=
7
8  Name: CC-PVDZ
9  Role: ORBITAL
10 Keyword: BASIS
11 atoms 1, 12-13 entry O line 190 file /Users/
    tatianefx/psi4conda/share/psi4/basis/cc-pvdz.gbs
12 atoms 2, 14 entry N line 160 file /Users/
    tatianefx/psi4conda/share/psi4/basis/cc-pvdz.gbs
13 atoms 3-5, 15-16 entry C line 130 file /Users/
    tatianefx/psi4conda/share/psi4/basis/cc-pvdz.gbs
14 atoms 6-11, 17-20 entry H line 20 file /Users/
    tatianefx/psi4conda/share/psi4/basis/cc-pvdz.gbs
15
16 There are an even number of electrons - assuming singlet.
17 Specify the multiplicity in the molecule input block.
18
19
20 -----
21                          SCF
22          by Justin Turney, Rob Parrish, and Andy Simmonett
23                          RHF Reference
24                  1 Threads, 476 MiB Core
25 -----
26
27 ==> Geometry <==
28
29 Molecular point group: c1
30 Full point group: C1
31
32 Geometry (in Angstrom), charge = 0, multiplicity = 1:
33
34      Center          X          Y          Z
35      Mass
36 -----
37
38      O          0.469665327601      1.352425119829
39          0.312281858816      15.994914619560
40      N          -2.190034672399      1.386625119829
41          0.899081858816      14.003074004780
42      C          -1.488034672399      0.122325119829

```

```

39      C      1.110681858816      12.000000000000
          -2.114034672399      -0.964074880171
          0.246781858816      12.000000000000
40      C      -0.032534672399      0.326225119829
          0.747781858816      12.000000000000
41      H      -1.546134672399      -0.137774880171
          2.172781858816      1.007825032070
42      H      -1.637534672399      -1.933574880171
          0.429181858816      1.007825032070
43      H      -3.181834672399      -1.076774880171
          0.464881858816      1.007825032070
44      H      -2.004334672399      -0.743774880171
          -0.821418141184      1.007825032070
45      H      -3.171134672399      1.277925119829
          1.154581858816      1.007825032070
46      H      -2.186634672399      1.623425119829
          -0.092918141184      1.007825032070
47      O      1.772065327601      -2.008674880171
          -0.277218141184      15.994914619560
48      O      -0.188034672399      -0.867574880171
          -0.411518141184      15.994914619560
49      N      0.991165327601      1.443525119829
          -1.061618141184      14.003074004780
50      C      1.869765327601      0.301125119829
          -0.869018141184      12.000000000000
51      C      1.031865327601      -0.895874880171
          -0.514018141184      12.000000000000
52      H      2.570965327601      0.506825119829
          -0.055618141184      1.007825032070
53      H      2.420965327601      0.096325119829
          -1.790818141184      1.007825032070
54      H      0.315165327601      1.243225119829
          -1.797618141184      1.007825032070
55      H      1.211765327601      -2.774874880171
          -0.029318141184      1.007825032070
56
57      Running in c1 symmetry.
58
59      Rotational constants: A =      0.07487      B =      0.04055      C =
          0.02972 [cm-1]
60      Rotational constants: A =      2244.52566      B =      1215.77934      C =
          890.83818 [MHz]
61      Nuclear repulsion =      606.712782605637358
62
63      Charge      = 0
64      Multiplicity = 1
65      Electrons   = 78

```

```

66  Nalpha          = 39
67  Nbeta           = 39
68
69  ==> Algorithm <==
70
71  SCF Algorithm Type is DF.
72  DIIS enabled.
73  MOM disabled.
74  Fractional occupation disabled.
75  Guess Type is SAD.
76  Energy threshold = 1.00e-06
77  Density threshold = 1.00e-06
78  Integral threshold = 0.00e+00
79
80  ==> Primary Basis <==
81
82  Basis Set: CC-PVDZ
83  Blend: CC-PVDZ
84  Number of shells: 90
85  Number of basis function: 190
86  Number of Cartesian functions: 200
87  Spherical Harmonics?: true
88  Max angular momentum: 2
89
90  => Loading Basis Set <=
91
92  Name: (CC-PVDZ AUX)
93  Role: JKFIT
94  Keyword: DF_BASIS_SCF
95  atoms 1, 12-13 entry O line 220 file /Users/
    tatianefx/psi4conda/share/psi4/basis/cc-pvdz-jkfit.gbs
96  atoms 2, 14 entry N line 170 file /Users/
    tatianefx/psi4conda/share/psi4/basis/cc-pvdz-jkfit.gbs
97  atoms 3-5, 15-16 entry C line 120 file /Users/
    tatianefx/psi4conda/share/psi4/basis/cc-pvdz-jkfit.gbs
98  atoms 6-11, 17-20 entry H line 50 file /Users/
    tatianefx/psi4conda/share/psi4/basis/cc-pvdz-jkfit.gbs
99
100 ==> Pre-Iterations <==
101
102 -----
103 Irrep   Nso     Nmo     Nalpha  Nbeta  Ndocc  Nsocc
104 -----
105 A        190     190        0        0        0        0
106 -----
107 Total    190     190     39     39     39        0
108 -----

```

```

109
110 ==> Integral Setup <==
111
112 ==> DFJK: Density-Fitted J/K Matrices <==
113
114     J tasked:                Yes
115     K tasked:                Yes
116     wK tasked:               No
117     OpenMP threads:          1
118     Integrals threads:        1
119     Memory (MB):              357
120     Algorithm:                Core
121     Integral Cache:           NONE
122     Schwarz Cutoff:           1E-12
123     Fitting Condition:        1E-12
124
125 => Auxiliary Basis Set <=
126
127 Basis Set: (CC-PVDZ AUX)
128 Blend: CC-PVDZ-JKFIT
129 Number of shells: 330
130 Number of basis function: 930
131 Number of Cartesian functions: 1060
132 Spherical Harmonics?: true
133 Max angular momentum: 3
134
135 Minimum eigenvalue in the overlap matrix is 1.8767339376E-03.
136 Using Symmetric Orthogonalization.
137
138 SCF Guess: Superposition of Atomic Densities via on-the-fly atomic
      UHF.
139
140 ==> Iterations <==
141
142                               Total Energy          Delta E          RMS |[F,P
                               ]|
143
144 @DF-RHF iter   0:  -533.54061727183250    -5.33541e+02    3.25189e
      -02
145 @DF-RHF iter   1:  -528.00911812942081     5.53150e+00    3.57812e
      -03
146 @DF-RHF iter   2:  -528.25567839698795    -2.46560e-01    1.86138e
      -03 DIIS
147 @DF-RHF iter   3:  -528.30986368160279    -5.41853e-02    5.23246e
      -04 DIIS
148 @DF-RHF iter   4:  -528.32003224868868    -1.01686e-02    2.09826e
      -04 DIIS

```

```

149   @DF-RHF iter    5:  -528.32217854696523   -2.14630e-03   9.73489e
      -05 DIIS
150   @DF-RHF iter    6:  -528.32311451905150   -9.35972e-04   6.05739e
      -05 DIIS
151   @DF-RHF iter    7:  -528.32353641182920   -4.21893e-04   3.31284e
      -05 DIIS
152   @DF-RHF iter    8:  -528.32372577622971   -1.89364e-04   1.61351e
      -05 DIIS
153   @DF-RHF iter    9:  -528.32377323680885   -4.74606e-05   6.50710e
      -06 DIIS
154   @DF-RHF iter   10:  -528.32377886243069   -5.62562e-06   2.90056e
      -06 DIIS
155   @DF-RHF iter   11:  -528.32377964076602   -7.78335e-07   1.23740e
      -06 DIIS
156   @DF-RHF iter   12:  -528.32377979516195   -1.54396e-07   4.26318e
      -07 DIIS
157
158   ==> Post-Iterations <==
159
160   Orbital Energies (a.u.)
161   -----
162
163   Doubly Occupied:
164
165       1A      -20.651757       2A      -20.596898       3A      -20.583467
166       4A      -15.653116       5A      -15.516712       6A      -11.446196
167       7A      -11.306631       8A      -11.248706       9A      -11.229938
168      10A      -11.194000      11A      -1.529343      12A      -1.469285
169      13A      -1.418115      14A      -1.206179      15A      -1.158292
170      16A      -1.017699      17A      -0.988951      18A      -0.876682
171      19A      -0.850070      20A      -0.791896      21A      -0.778994
172      22A      -0.753486      23A      -0.728469      24A      -0.666070
173      25A      -0.651983      26A      -0.631440      27A      -0.606800
174      28A      -0.595029      29A      -0.586490      30A      -0.568347
175      31A      -0.556577      32A      -0.534709      33A      -0.515581
176      34A      -0.494839      35A      -0.479390      36A      -0.452007
177      37A      -0.448204      38A      -0.365318      39A      -0.229218
178
179   Virtual:
180
181      40A       0.125223      41A       0.142840      42A       0.161885
182      43A       0.197619      44A       0.206312      45A       0.221827
183      46A       0.250645      47A       0.260075      48A       0.263785
184      49A       0.287565      50A       0.298686      51A       0.319904
185      52A       0.337526      53A       0.357802      54A       0.408050
186      55A       0.431685      56A       0.438975      57A       0.452801
187      58A       0.512037      59A       0.538313      60A       0.585788

```

188	61A	0.638310	62A	0.653164	63A	0.666632
189	64A	0.689624	65A	0.700580	66A	0.715276
190	67A	0.723628	68A	0.735767	69A	0.757730
191	70A	0.775341	71A	0.809209	72A	0.820838
192	73A	0.842135	74A	0.843690	75A	0.857445
193	76A	0.865420	77A	0.874326	78A	0.901197
194	79A	0.914333	80A	0.915335	81A	0.930348
195	82A	0.941026	83A	0.946505	84A	0.976694
196	85A	0.998162	86A	1.011478	87A	1.045281
197	88A	1.046918	89A	1.064866	90A	1.083998
198	91A	1.111849	92A	1.145594	93A	1.165090
199	94A	1.186099	95A	1.195909	96A	1.206784
200	97A	1.242409	98A	1.280994	99A	1.311702
201	100A	1.337177	101A	1.352528	102A	1.370438
202	103A	1.386534	104A	1.432435	105A	1.484624
203	106A	1.497420	107A	1.519316	108A	1.536736
204	109A	1.559236	110A	1.593629	111A	1.606132
205	112A	1.620473	113A	1.632828	114A	1.635787
206	115A	1.678752	116A	1.701215	117A	1.750897
207	118A	1.770998	119A	1.790315	120A	1.823278
208	121A	1.857499	122A	1.881797	123A	1.884430
209	124A	1.894787	125A	1.902542	126A	1.922912
210	127A	1.940313	128A	1.951659	129A	1.965389
211	130A	1.973800	131A	1.993035	132A	2.007791
212	133A	2.026074	134A	2.034663	135A	2.061889
213	136A	2.083289	137A	2.097789	138A	2.144060
214	139A	2.153020	140A	2.165679	141A	2.171982
215	142A	2.226995	143A	2.247218	144A	2.256733
216	145A	2.296705	146A	2.310719	147A	2.339128
217	148A	2.361784	149A	2.370685	150A	2.410102
218	151A	2.416419	152A	2.421380	153A	2.441824
219	154A	2.489561	155A	2.542825	156A	2.571045
220	157A	2.577635	158A	2.589245	159A	2.615163
221	160A	2.658924	161A	2.700617	162A	2.708384
222	163A	2.779545	164A	2.787644	165A	2.819008
223	166A	2.843379	167A	2.870168	168A	2.887230
224	169A	2.918238	170A	2.982293	171A	3.004436
225	172A	3.056454	173A	3.111051	174A	3.178622
226	175A	3.208370	176A	3.230784	177A	3.233736
227	178A	3.310014	179A	3.365354	180A	3.426525
228	181A	3.458411	182A	3.479476	183A	3.484834
229	184A	3.554891	185A	3.631752	186A	3.817471
230	187A	3.866359	188A	3.897391	189A	3.948736
231	190A	4.088810				
232						
233	Final Occupation by Irrep:					
234	A					

```
235      DOCC [    39 ]
236
237      Energy converged.
238
239      @DF-RHF Final Energy:   -528.32377979516195
240
241      => Energetics <=
242
243      Nuclear Repulsion Energy =                606.7127826056373578
244      One-Electron Energy =              -1927.3193335674816353
245      Two-Electron Energy =                792.2827711666823234
246      DFT Exchange-Correlation Energy =          0.0000000000000000
247      Empirical Dispersion Energy =          0.0000000000000000
248      PCM Polarization Energy =          0.0000000000000000
249      EFP Energy =                0.0000000000000000
250      Total Energy =              -528.3237797951619541
251
252
253
254      Properties will be evaluated at    0.000000,    0.000000,    0.000000
          Bohr
255
256      Properties computed using the SCF density matrix
257
258      Nuclear Dipole Moment: (a.u.)
259      X:    -6.7446      Y:    -1.8146      Z:    -0.3447
260
261      Electronic Dipole Moment: (a.u.)
262      X:     7.1380      Y:     0.3986      Z:    -1.2883
263
264      Dipole Moment: (a.u.)
265      X:     0.3934      Y:    -1.4160      Z:    -1.6330      Total:
          2.1969
266
267      Dipole Moment: (Debye)
268      X:     0.9999      Y:    -3.5990      Z:    -4.1507      Total:
          5.5840
269
270
271      *** tstop() called on Tatianes-MacBook-Air.local at Thu Jun 20
          14:03:20 2019
272      Module time:
273      user time   =          20.10 seconds =          0.33 minutes
274      system time =           0.55 seconds =          0.01 minutes
275      total time  =           16 seconds =          0.27 minutes
276      Total time:
277      user time   =          798.70 seconds =         13.31 minutes
```

```

278 system time =          22.28 seconds =          0.37 minutes
279 total time  =          624 seconds =         10.40 minutes

```

Listagem C.3 – timer.dat

```

1 Host: Tatianes-MacBook-Air.local
2
3 Timers On : Thu Jun 20 13:52:50 2019
4 Timers Off: Thu Jun 20 14:09:30 2019
5
6 Wall Time:      1000.00 seconds
7
8 JK: (A|mn) :      351.28u      0.75s      212.10w      200 calls
9 JK: (A|Q)^-1/2:      120.57u      2.68s      60.17w      200 calls
10 JK: (Q|mn) :      143.38u      0.73s      23.67w      200 calls
11 HF: Form H :        0.05u      0.17s      0.173638w      40 calls
12 HF: Form S/X:       1.08u      0.02s      0.631288w      40 calls
13 HF: Guess :        8.57u      0.12s      5.266284w      40 calls
14 SAD Guess :        8.45u      0.12s      5.159809w      40 calls
15 JK: D :          0.20u      0.03s      0.125210w     1520 calls
16 JK: US02A0 :        0.00u      0.00s      0.034458w     1520 calls
17 JK: JK :        797.10u     26.95s      637.17w     1520 calls
18 JK: J :         33.15u      0.23s      10.34w     1520 calls
19 JK: J1 :        16.82u      0.12s      5.232191w     2400 calls
20 JK: J2 :        16.15u      0.10s      4.982653w     2400 calls
21 JK: K :        754.18u     16.55s      615.17w     1520 calls
22 JK: K1 :        640.93u     16.20s      580.47w     2320 calls
23 JK: K2 :        113.20u      0.35s      34.67w     2320 calls
24 JK: A02US0 :        0.00u      0.00s      0.003373w     1520 calls
25 DIISManager::add_entry:      0.07u      0.20s      0.278112w     1480
    calls
26 DIISManager::extrapolate:      0.68u      1.00s      0.977797w     1440
    calls
27 DIISManager::extrapolate: bMatrix setup:      0.18u      0.55s
    0.488405w     1440 calls
28 DIISManager::extrapolate: bMatrix pseudoinverse:      0.13u
    0.00s      0.035510w     1440 calls
29 DIISManager::extrapolate: form new data:      0.37u      0.45s
    0.438143w     1440 calls
30 HF: Form G :      797.38u     27.03s      637.37w      640 calls
31 HF: Form F :        0.00u      0.00s      0.045055w      640 calls
32 HF: DIIS :         3.82u      1.38s      2.394890w      640 calls
33 HF: Form C :       16.47u      0.17s      9.445859w      640 calls
34 HF: Form D :        0.17u      0.02s      0.114420w      640 calls
35
36 *****

```

C.2 Ala-Gly-Ser

Listagem C.4 – Log

```
1 Enter an amino acid sequence: AGS
2
3 Generates Population
4
5 Fitness: -846.3830882731716
6 Fitness: -846.0706046134601
7 Fitness: -839.2166319125993
8 Fitness: -848.2767413591066
9 Fitness: -848.4628586178002
10
11 GENERATION 0
12 Fitness: -848.0544173576209
13 Fitness: -848.5715760591886
14 Fitness: -847.4228379747836
15 Fitness: -848.0889970162556
16 Fitness: -848.5030199849069
17 Fitness: -848.7907306374677
18 Fitness: -847.9813451674539
19
20 GENERATION 1
21 Fitness: -848.4294580634481
22 Fitness: -848.5417432377287
23 Fitness: -847.3553365558872
24 Fitness: -847.8756690835153
25 Fitness: -848.7304757529168
26 Fitness: -848.5875372267864
27 Fitness: -848.7896508548256
28
29 GENERATION 2
30 Fitness: -848.7642078021361
31 Fitness: -848.3379051435824
32 Fitness: -847.8172532959155
33 Fitness: -848.7516754683762
34 Fitness: -848.0128849165753
35 Fitness: -847.9618665998814
36 Fitness: -848.6445946509956
37
38 GENERATION 3
39 Fitness: -848.7352089781205
40 Fitness: -847.9835169523552
41 Fitness: -848.0481310278327
42 Fitness: -848.4686272912074
43 Fitness: -848.3916479171955
44 Fitness: -848.6940267231785
```

```
45 Fitness: -848.7755377064457
46
47 GENERATION 4
48 Fitness: -848.5738255118922
49 Fitness: -847.9527506000068
50 Fitness: -848.480961256966
51 Fitness: -847.3987093248422
52 Fitness: -848.4477822678394
53 Fitness: -848.7141484055456
54 Fitness: -848.7900243556691
55
56
57 The best fitness: -848.7900243556691
58
59 The best geometry:
60
61 O 1.2492 1.1165 -0.4047
62 N -1.4105 1.1507 0.1821
63 C -0.7085 -0.1136 0.3937
64 C -1.3345 -1.2000 -0.4702
65 C 0.7470 0.0903 0.0308
66 H -0.7666 -0.3737 1.4558
67 H -0.8580 -2.1695 -0.2878
68 H -2.4023 -1.3127 -0.2521
69 H -1.2248 -0.9797 -1.5384
70 H -2.3916 1.0420 0.4376
71 H -1.4071 1.3875 -0.8099
72 O 0.9598 -1.3807 -0.1162
73 N 0.5750 -0.6240 -2.6527
74 C 1.9543 -0.8319 -2.2422
75 C 1.9685 -1.2512 -0.7982
76 H 2.5179 0.0983 -2.3551
77 H 2.4093 -1.6175 -2.8514
78 H 0.0386 -1.4802 -2.5186
79 O 1.1245 4.7820 3.8308
80 O 0.3416 1.9398 4.4038
81 O -0.4194 1.6761 2.2792
82 N 2.1655 2.2177 1.3911
83 C 1.7394 2.7571 2.6815
84 C 1.5343 4.2652 2.5691
85 C 0.4410 2.0678 3.0546
86 H 2.4975 2.5178 3.4357
87 H 2.4558 4.7760 2.2697
88 H 0.7519 4.5137 1.8437
89 H 2.2579 1.2046 1.4510
90 H 1.8283 4.5818 4.4712
91 H -0.4940 1.4991 4.6685
```

```

92 THE END
93
94 Process finished with exit code 0

```

Listagem C.5 – output.dat

```

1  Memory set to 476.837 MiB by Python driver.
2
3  *** tstart() called on Tatianes-MacBook-Air.local
4  *** at Thu Jun 20 15:10:40 2019
5
6  ==> Loading Basis Set <=
7
8  Name: CC-PVDZ
9  Role: ORBITAL
10 Keyword: BASIS
11 atoms 1, 12, 19-21      entry 0          line 190 file /Users/
    tatianefx/psi4conda/share/psi4/basis/cc-pvdz.gbs
12 atoms 2, 13, 22        entry N          line 160 file /Users/
    tatianefx/psi4conda/share/psi4/basis/cc-pvdz.gbs
13 atoms 3-5, 14-15, 23-25 entry C          line 130 file /Users/
    tatianefx/psi4conda/share/psi4/basis/cc-pvdz.gbs
14 atoms 6-11, 16-18, 26-31 entry H          line 20 file /Users/
    tatianefx/psi4conda/share/psi4/basis/cc-pvdz.gbs
15
16 There are an even number of electrons - assuming singlet.
17 Specify the multiplicity in the molecule input block.
18
19
20 -----
21                      SCF
22      by Justin Turney, Rob Parrish, and Andy Simmonett
23                      RHF Reference
24                      1 Threads,      476 MiB Core
25 -----
26
27 ==> Geometry <==
28
29 Molecular point group: c1
30 Full point group: C1
31
32 Geometry (in Angstrom), charge = 0, multiplicity = 1:
33
34      Center          X          Y          Z
35      Mass
36 -----
37 -----

```

36	O	0.596906740839	0.036596034586
		-1.334948760147	15.994914619560
37	N	-2.062793259161	0.070796034586
		-0.748148760147	14.003074004780
38	C	-1.360793259161	-1.193503965414
		-0.536548760147	12.000000000000
39	C	-1.986793259161	-2.279903965414
		-1.400448760147	12.000000000000
40	C	0.094706740839	-0.989603965414
		-0.899448760147	12.000000000000
41	H	-1.418893259161	-1.453603965414
		0.525551239853	1.007825032070
42	H	-1.510293259161	-3.249403965414
		-1.218048760147	1.007825032070
43	H	-3.054593259161	-2.392603965414
		-1.182348760147	1.007825032070
44	H	-1.877093259161	-2.059603965414
		-2.468648760147	1.007825032070
45	H	-3.043893259161	-0.037903965414
		-0.492648760147	1.007825032070
46	H	-2.059393259161	0.307596034586
		-1.740148760147	1.007825032070
47	O	0.307506740839	-2.460603965414
		-1.046448760147	15.994914619560
48	N	-0.077293259161	-1.703903965414
		-3.582948760147	14.003074004780
49	C	1.302006740839	-1.911803965414
		-3.172448760147	12.000000000000
50	C	1.316206740839	-2.331103965414
		-1.728448760147	12.000000000000
51	H	1.865606740839	-0.981603965414
		-3.285348760147	1.007825032070
52	H	1.757006740839	-2.697403965414
		-3.781648760147	1.007825032070
53	H	-0.613693259161	-2.560103965414
		-3.448848760147	1.007825032070
54	O	0.472206740839	3.702096034586
		2.900551239853	15.994914619560
55	O	-0.310693259161	0.859896034586
		3.473551239853	15.994914619560
56	O	-1.071693259161	0.596196034586
		1.348951239853	15.994914619560
57	N	1.513206740839	1.137796034586
		0.460851239853	14.003074004780
58	C	1.087106740839	1.677196034586
		1.751251239853	12.000000000000
59	C	0.882006740839	3.185296034586

```

        1.638851239853      12.000000000000
60      C      -0.211293259161      0.987896034586
        2.124351239853      12.000000000000
61      H      1.845206740839      1.437896034586
        2.505451239853      1.007825032070
62      H      1.803506740839      3.696096034586
        1.339451239853      1.007825032070
63      H      0.099606740839      3.433796034586
        0.913451239853      1.007825032070
64      H      1.605606740839      0.124696034586
        0.520751239853      1.007825032070
65      H      1.176006740839      3.501896034586
        3.540951239853      1.007825032070
66      H      -1.146293259161      0.419196034586
        3.738251239853      1.007825032070
67
68  Running in c1 symmetry.
69
70  Rotational constants: A =      0.03440  B =      0.00936  C =
      0.00847 [cm-1]
71  Rotational constants: A =    1031.22492  B =      280.66287  C =
      253.81490 [MHz]
72  Nuclear repulsion = 1208.145979741552310
73
74  Charge      = 0
75  Multiplicity = 1
76  Electrons   = 124
77  Nalpha      = 62
78  Nbeta       = 62
79
80  ==> Algorithm <==
81
82  SCF Algorithm Type is DF.
83  DIIS enabled.
84  MOM disabled.
85  Fractional occupation disabled.
86  Guess Type is SAD.
87  Energy threshold  = 1.00e-06
88  Density threshold = 1.00e-06
89  Integral threshold = 0.00e+00
90
91  ==> Primary Basis <==
92
93  Basis Set: CC-PVDZ
94  Blend: CC-PVDZ
95  Number of shells: 141
96  Number of basis function: 299

```

```

97   Number of Cartesian functions: 315
98   Spherical Harmonics?: true
99   Max angular momentum: 2
100
101  => Loading Basis Set <=
102
103   Name: (CC-PVDZ AUX)
104   Role: JKFIT
105   Keyword: DF_BASIS_SCF
106   atoms 1, 12, 19-21      entry 0          line 220 file /Users/
        tatianefx/psi4conda/share/psi4/basis/cc-pvdz-jkfit.gbs
107   atoms 2, 13, 22        entry N          line 170 file /Users/
        tatianefx/psi4conda/share/psi4/basis/cc-pvdz-jkfit.gbs
108   atoms 3-5, 14-15, 23-25 entry C          line 120 file /Users/
        tatianefx/psi4conda/share/psi4/basis/cc-pvdz-jkfit.gbs
109   atoms 6-11, 16-18, 26-31 entry H          line 50 file /Users/
        tatianefx/psi4conda/share/psi4/basis/cc-pvdz-jkfit.gbs
110
111  ==> Pre-Iterations <==
112
113  -----
114   Irrep   Nso      Nmo      Nalpha   Nbeta   Ndocc   Nsocc
115  -----
116   A        299      299         0        0         0         0
117  -----
118   Total     299      299        62        62        62         0
119  -----
120
121  ==> Integral Setup <==
122
123  ==> DFJK: Density-Fitted J/K Matrices <==
124
125   J tasked:                Yes
126   K tasked:                Yes
127   wK tasked:               No
128   OpenMP threads:          1
129   Integrals threads:        1
130   Memory (MB):             357
131   Algorithm:                Disk
132   Integral Cache:           NONE
133   Schwarz Cutoff:           1E-12
134   Fitting Condition:        1E-12
135
136  => Auxiliary Basis Set <=
137
138   Basis Set: (CC-PVDZ AUX)
139   Blend: CC-PVDZ-JKFIT

```

```

140     Number of shells: 519
141     Number of basis function: 1465
142     Number of Cartesian functions: 1671
143     Spherical Harmonics?: true
144     Max angular momentum: 3
145
146     Minimum eigenvalue in the overlap matrix is 2.1022888044E-03.
147     Using Symmetric Orthogonalization.
148
149     SCF Guess: Superposition of Atomic Densities via on-the-fly atomic
        UHF.
150
151     ==> Iterations <==
152
153                                     Total Energy          Delta E          RMS |[F,P
                                                ]|
154
155     @DF-RHF iter    0:   -854.34626729106083   -8.54346e+02   2.37108e
        -02
156     @DF-RHF iter    1:   -847.81759515928195    6.52867e+00   3.99879e
        -03
157     @DF-RHF iter    2:   -843.59030907205442    4.22729e+00   8.25865e
        -03 DIIS
158     @DF-RHF iter    3:   -848.40657278539231   -4.81626e+00   2.56848e
        -03 DIIS
159     @DF-RHF iter    4:   -848.64797758006068   -2.41405e-01   1.12694e
        -03 DIIS
160     @DF-RHF iter    5:   -848.73486452160080   -8.68869e-02   5.08480e
        -04 DIIS
161     @DF-RHF iter    6:   -848.75909310825364   -2.42286e-02   2.70194e
        -04 DIIS
162     @DF-RHF iter    7:   -848.76217913627158   -3.08603e-03   5.24500e
        -04 DIIS
163     @DF-RHF iter    8:   -848.76443789872565   -2.25876e-03   5.07575e
        -04 DIIS
164     @DF-RHF iter    9:   -848.76811526739971   -3.67737e-03   4.96730e
        -04 DIIS
165     @DF-RHF iter   10:   -848.77647410861721   -8.35884e-03   3.98385e
        -04 DIIS
166     @DF-RHF iter   11:   -848.78342535775880   -6.95125e-03   2.77003e
        -04 DIIS
167     @DF-RHF iter   12:   -848.78806691013654   -4.64155e-03   9.16556e
        -05 DIIS
168     @DF-RHF iter   13:   -848.78849650439292   -4.29594e-04   1.37360e
        -04 DIIS
169     @DF-RHF iter   14:   -848.78935236215875   -8.55858e-04   8.37474e
        -05 DIIS

```



```

170   @DF-RHF iter 15:  -848.78979694419002  -4.44582e-04  4.72388e
      -05 DIIS
171   @DF-RHF iter 16:  -848.78992942406421  -1.32480e-04  3.82842e
      -05 DIIS
172   @DF-RHF iter 17:  -848.78997840933505  -4.89853e-05  1.52494e
      -05 DIIS
173   @DF-RHF iter 18:  -848.79000970081051  -3.12915e-05  6.43972e
      -06 DIIS
174   @DF-RHF iter 19:  -848.79001935877181  -9.65796e-06  3.04995e
      -06 DIIS
175   @DF-RHF iter 20:  -848.79002259432343  -3.23555e-06  2.47646e
      -06 DIIS
176   @DF-RHF iter 21:  -848.79002404903679  -1.45471e-06  1.08091e
      -06 DIIS
177   @DF-RHF iter 22:  -848.79002435566906  -3.06632e-07  9.84833e
      -07 DIIS
178
179 ==> Post-Iterations <==
180
181   Orbital Energies (a.u.)
182   -----
183
184   Doubly Occupied:
185
186       1A    -20.684139      2A    -20.639556      3A    -20.605307
187       4A    -20.547183      5A    -20.540613      6A    -15.587449
188       7A    -15.558788      8A    -15.526817      9A    -11.440179
189      10A    -11.381538     11A    -11.373779     12A    -11.290913
190      13A    -11.283139     14A    -11.275363     15A    -11.268248
191      16A    -11.229514     17A     -1.574789     18A     -1.447574
192      19A     -1.444756     20A     -1.349715     21A     -1.346896
193      22A     -1.190474     23A     -1.177056     24A     -1.151072
194      25A     -1.045703     26A     -0.968613     27A     -0.953340
195      28A     -0.935268     29A     -0.889489     30A     -0.832150
196      31A     -0.811730     32A     -0.781564     33A     -0.749557
197      34A     -0.728621     35A     -0.707404     36A     -0.695873
198      37A     -0.677905     38A     -0.675583     39A     -0.648474
199      40A     -0.641117     41A     -0.631028     42A     -0.625607
200      43A     -0.614472     44A     -0.595916     45A     -0.582216
201      46A     -0.577241     47A     -0.572187     48A     -0.560137
202      49A     -0.554598     50A     -0.544044     51A     -0.535150
203      52A     -0.512502     53A     -0.484323     54A     -0.479649
204      55A     -0.470689     56A     -0.453104     57A     -0.449402
205      58A     -0.434035     59A     -0.405656     60A     -0.370141
206      61A     -0.336483     62A     -0.209253
207
208   Virtual:

```

209						
210	63A	-0.082202	64A	0.072540	65A	0.105681
211	66A	0.167519	67A	0.174196	68A	0.186118
212	69A	0.200736	70A	0.204992	71A	0.219088
213	72A	0.223029	73A	0.240894	74A	0.242691
214	75A	0.250084	76A	0.260325	77A	0.274518
215	78A	0.279083	79A	0.286715	80A	0.290233
216	81A	0.311709	82A	0.329955	83A	0.351125
217	84A	0.381502	85A	0.396052	86A	0.401436
218	87A	0.410026	88A	0.424726	89A	0.435208
219	90A	0.462200	91A	0.487854	92A	0.510578
220	93A	0.520472	94A	0.552775	95A	0.558900
221	96A	0.616144	97A	0.641492	98A	0.655189
222	99A	0.664515	100A	0.679985	101A	0.692455
223	102A	0.698708	103A	0.702466	104A	0.712918
224	105A	0.721785	106A	0.730316	107A	0.739386
225	108A	0.745090	109A	0.759686	110A	0.774328
226	111A	0.781573	112A	0.788652	113A	0.793846
227	114A	0.817979	115A	0.826386	116A	0.836650
228	117A	0.838885	118A	0.846134	119A	0.850653
229	120A	0.854664	121A	0.866609	122A	0.874732
230	123A	0.878303	124A	0.881970	125A	0.887510
231	126A	0.898621	127A	0.905827	128A	0.910325
232	129A	0.926642	130A	0.928268	131A	0.932999
233	132A	0.957671	133A	0.978702	134A	1.002990
234	135A	1.008486	136A	1.026895	137A	1.031603
235	138A	1.054965	139A	1.064960	140A	1.087506
236	141A	1.097111	142A	1.116645	143A	1.121442
237	144A	1.149761	145A	1.157507	146A	1.166537
238	147A	1.183722	148A	1.189032	149A	1.207263
239	150A	1.219608	151A	1.230519	152A	1.250287
240	153A	1.259329	154A	1.272248	155A	1.293650
241	156A	1.296944	157A	1.302585	158A	1.322957
242	159A	1.334791	160A	1.347069	161A	1.361597
243	162A	1.378009	163A	1.395626	164A	1.415763
244	165A	1.420590	166A	1.431139	167A	1.443124
245	168A	1.464600	169A	1.480514	170A	1.492259
246	171A	1.500298	172A	1.525757	173A	1.530902
247	174A	1.551513	175A	1.578809	176A	1.601395
248	177A	1.610499	178A	1.628768	179A	1.631840
249	180A	1.641430	181A	1.666885	182A	1.669657
250	183A	1.697601	184A	1.714693	185A	1.729426
251	186A	1.772070	187A	1.782516	188A	1.783992
252	189A	1.798902	190A	1.801374	191A	1.815014
253	192A	1.831701	193A	1.846430	194A	1.853306
254	195A	1.875423	196A	1.882666	197A	1.896862
255	198A	1.902560	199A	1.925818	200A	1.929869

256	201A	1.933295	202A	1.946746	203A	1.948277
257	204A	1.960755	205A	1.969864	206A	1.983493
258	207A	1.989966	208A	2.003290	209A	2.009621
259	210A	2.013477	211A	2.020991	212A	2.037531
260	213A	2.049824	214A	2.058620	215A	2.067519
261	216A	2.080405	217A	2.096766	218A	2.113677
262	219A	2.131242	220A	2.138188	221A	2.151339
263	222A	2.157301	223A	2.182918	224A	2.184933
264	225A	2.197307	226A	2.220168	227A	2.243645
265	228A	2.271548	229A	2.291100	230A	2.303046
266	231A	2.326541	232A	2.335286	233A	2.353899
267	234A	2.363791	235A	2.379725	236A	2.390717
268	237A	2.409673	238A	2.431220	239A	2.441598
269	240A	2.464785	241A	2.475696	242A	2.498455
270	243A	2.512894	244A	2.534368	245A	2.562776
271	246A	2.569459	247A	2.591480	248A	2.606432
272	249A	2.616240	250A	2.634025	251A	2.658403
273	252A	2.675441	253A	2.703498	254A	2.712331
274	255A	2.742209	256A	2.777695	257A	2.784589
275	258A	2.793539	259A	2.812332	260A	2.818947
276	261A	2.842105	262A	2.851785	263A	2.861123
277	264A	2.876513	265A	2.899744	266A	2.930248
278	267A	2.960915	268A	2.962696	269A	2.975990
279	270A	2.986928	271A	3.013126	272A	3.054497
280	273A	3.077354	274A	3.095321	275A	3.108892
281	276A	3.211778	277A	3.226412	278A	3.284701
282	279A	3.322929	280A	3.329245	281A	3.333523
283	282A	3.354226	283A	3.408144	284A	3.430993
284	285A	3.443377	286A	3.458920	287A	3.466055
285	288A	3.473636	289A	3.552718	290A	3.569361
286	291A	3.577116	292A	3.655876	293A	3.657170
287	294A	3.859448	295A	3.876518	296A	3.945231
288	297A	3.951347	298A	4.096823	299A	4.108175

289

290 Final Occupation by Irrep:

291 A

292 DOCC [62]

293

294 Energy converged.

295

296 @DF-RHF Final Energy: -848.79002435566906

297

298 => Energetics <=

299

300 Nuclear Repulsion Energy = 1208.1459797415523099

301 One-Electron Energy = -3564.5620784896968871

302 Two-Electron Energy = 1507.6260743924756298

```

303     DFT Exchange-Correlation Energy =          0.000000000000000000
304     Empirical Dispersion Energy =          0.000000000000000000
305     PCM Polarization Energy =          0.000000000000000000
306     EFP Energy =          0.000000000000000000
307     Total Energy =          -848.7900243556689475
308
309
310
311 Properties will be evaluated at    0.000000,    0.000000,    0.000000
      Bohr
312
313 Properties computed using the SCF density matrix
314
315     Nuclear Dipole Moment: (a.u.)
316         X:      -4.2839          Y:      -2.3395          Z:      -4.2134
317
318     Electronic Dipole Moment: (a.u.)
319         X:       4.1341          Y:       0.3065          Z:       4.5563
320
321     Dipole Moment: (a.u.)
322         X:      -0.1498          Y:      -2.0329          Z:       0.3429      Total:
      2.0671
323
324     Dipole Moment: (Debye)
325         X:      -0.3808          Y:      -5.1672          Z:       0.8715      Total:
      5.2540
326
327
328 *** tstop() called on Tatianes-MacBook-Air.local at Thu Jun 20
      15:11:55 2019
329 Module time:
330     user time   =      114.54 seconds =      1.91 minutes
331     system time =       7.64 seconds =      0.13 minutes
332     total time  =       75 seconds =      1.25 minutes
333 Total time:
334     user time   =     4012.23 seconds =     66.87 minutes
335     system time =     258.48 seconds =      4.31 minutes
336     total time  =     3342 seconds =     55.70 minutes

```

Listagem C.6 – timer.dat

```

1 Host: Tatianes-MacBook-Air.local
2
3 Timers On : Thu Jun 20 14:16:07 2019
4 Timers Off: Thu Jun 20 15:11:55 2019
5
6 Wall Time:      3348.00 seconds
7

```

```

 8 JK: (A|Q)^-1:      517.85u      34.83s      249.00w      40 calls
 9 JK: (A|mn) :      473.10u      1.20s      279.68w      240 calls
10 JK: (Q|mn) :      762.72u      2.97s      556.53w      240 calls
11 JK: (Q|mn) Write:      0.57u      10.65s      9.330220w      80 calls
12 HF: Form H :      0.22u      0.35s      0.682354w      40 calls
13 HF: Form S/X:      4.22u      0.23s      2.444666w      40 calls
14 HF: Guess :      9.10u      0.13s      5.442978w      40 calls
15 SAD Guess :      8.95u      0.12s      5.320051w      40 calls
16 JK: (A|Q)^-1/2:      2.35u      0.02s      1.434192w      160 calls
17 JK: D :      0.68u      0.05s      0.308862w      1624 calls
18 JK: US02A0 :      0.18u      0.00s      0.088586w      1624 calls
19 JK: JK :      4675.28u      353.30s      1939.63w      1624 calls
20 JK: J :      128.52u      0.65s      40.30w      2408 calls
21 JK: J1 :      65.25u      0.37s      20.36w      3288 calls
22 JK: J2 :      62.37u      0.27s      19.29w      3288 calls
23 JK: K :      4427.67u      65.93s      1653.10w      2408 calls
24 JK: K1 :      2769.98u      61.13s      1214.68w      3208 calls
25 JK: K2 :      1657.53u      4.80s      438.30w      3208 calls
26 JK: A02US0 :      0.02u      0.00s      0.003713w      1624 calls
27 DIISManager::add_entry:      0.35u      0.78s      0.674346w      1584
    calls
28 DIISManager::extrapolate:      1.80u      3.00s      2.897846w      1544
    calls
29 DIISManager::extrapolate: bMatrix setup:      0.68u      1.78s
    1.559466w      1544 calls
30 DIISManager::extrapolate: bMatrix pseudoinverse:      0.03u
    0.00s      0.043117w      1544 calls
31 DIISManager::extrapolate: form new data:      1.07u      1.22s
    1.276245w      1544 calls
32 HF: Form G :      4676.73u      353.55s      1940.31w      744 calls
33 JK: (Q|mn) Read:      5.22u      144.92s      90.96w      1528 calls
34 HF: Form F :      0.18u      0.00s      0.140554w      744 calls
35 HF: DIIS :      16.95u      4.18s      7.824059w      744 calls
36 HF: Form C :      74.40u      2.58s      41.08w      744 calls
37 HF: Form D :      0.62u      0.07s      0.276748w      744 calls
38
39 *****

```

C.3 Ala-Gly-Ser-Glu

Listagem C.7 – Log

```
1 Enter an amino acid sequence: AGSE
2
3 Generates Population
4
5 Fitness: -1314.5322358300832
6 Fitness: -1318.2445169216492
7 Fitness: -1319.4036808054686
8 Fitness: -1320.2491615184488
9 Fitness: -1311.9971133079182
10
11 GENERATION 0
12 Fitness: -1314.4774096124097
13 Fitness: -1319.1729666265478
14 Fitness: -1318.256819441998
15 Fitness: -1319.35256456108
16 Fitness: 0
17 Fitness: -1320.5842688474065
18 Fitness: -1319.0596645647854
19
20 GENERATION 1
21 Fitness: -1314.0133361925934
22 Fitness: -1300.6473461830562
23 Fitness: -1293.3423883634364
24 Fitness: -1314.6747056647775
25 Fitness: -1316.7277528688442
26 Fitness: -1312.7067370592856
27 Fitness: -1315.1593021783635
28
29 GENERATION 2
30 Fitness: -1319.8111715352545
31 Fitness: -1320.1979365139189
32 Fitness: -1320.092123197193
33 Fitness: -1320.4350487075149
34 Fitness: -1319.8067951307096
35 Fitness: -1319.97929252328
36 Fitness: -1315.5373725715153
37
38 GENERATION 3
39 Fitness: -1319.4737511483606
40 Fitness: -1318.8879570034928
41 Fitness: -1314.9886462399982
42 Fitness: -1316.4801925602019
43 Fitness: -1315.8942698662195
44 Fitness: -1299.8282371017074
```

```
45 Fitness: -1319.508184138544
46
47 GENERATION 4
48 Fitness: -1318.5523473344301
49 Fitness: 0
50 Fitness: -1319.7261592897798
51 Fitness: -1320.0956240984865
52 Fitness: -1319.6743915695727
53 Fitness: -1296.0534312912594
54 Fitness: -1318.1876367722057
55
56
57 The best fitness: -1320.0956240984865
58
59 The best geometry:
60
61 O 1.2492 1.1165 -0.4047
62 N -1.4105 1.1507 0.1821
63 C -0.7085 -0.1136 0.3937
64 C -1.3345 -1.2000 -0.4702
65 C 0.7470 0.0903 0.0308
66 H -0.7666 -0.3737 1.4558
67 H -0.8580 -2.1695 -0.2878
68 H -2.4023 -1.3127 -0.2521
69 H -1.2248 -0.9797 -1.5384
70 H -2.3916 1.0420 0.4376
71 H -1.4071 1.3875 -0.8099
72 O 0.9591 -1.3813 -0.1156
73 N 1.2432 -0.5861 2.4224
74 C 0.8527 -1.9719 2.2189
75 C 0.7461 -2.2297 0.7413
76 H -0.1168 -2.1558 2.6898
77 H 1.6060 -2.6375 2.6488
78 H 2.1313 -0.4024 1.9574
79 O -4.6148 0.1384 -1.3991
80 O -2.3600 -0.0394 -1.6152
81 N -2.1583 2.5518 -0.6110
82 C -3.3898 2.1575 -1.2930
83 C -3.4678 2.8407 -2.6554
84 C -3.3569 0.6496 -1.4527
85 H -4.2469 2.4291 -0.6665
86 H -3.4612 3.9318 -2.5600
87 H -2.6345 2.5469 -3.3031
88 H -2.0808 2.0609 0.2788
89 H -4.6177 -0.8370 -1.5042
90 O 6.6398 -0.0485 -0.3901
91 O 2.0139 -2.3135 -2.1904
```

```

92 O  6.4116  1.3200 -2.1831
93 O  1.9072 -1.8507  0.0315
94 N  4.2140  2.5375 -1.1600
95 C  3.7360  0.3147 -0.1453
96 C  4.7670  1.4310 -0.3818
97 C  3.2753 -0.4540 -1.3867
98 C  6.0062  0.9238 -1.0979
99 C  2.3395 -1.5977 -1.0856
100 H  2.8572  0.7381  0.3598
101 H  4.1618 -0.4050  0.5671
102 H  5.0896  1.8336  0.5858
103 H  2.7585  0.2268 -2.0720
104 H  4.1437 -0.8774 -1.9034
105 H  3.3859  2.9040 -0.6923
106 H  7.4387 -0.3769 -0.8551
107 H  1.4005 -3.0502 -1.9820
108 THE END
109
110 Process finished with exit code 0

```

Listagem C.8 – output.dat

```

1  Memory set to 476.837 MiB by Python driver.
2
3  *** tstart() called on Tatianes-MacBook-Air.local
4  *** at Thu Jun 20 21:35:06 2019
5
6  => Loading Basis Set <=
7
8  Name: CC-PVDZ
9  Role: ORBITAL
10 Keyword: BASIS
11 atoms 1, 12, 19-20, 30-33      entry O          line 190 file
    /Users/tatianefx/psi4conda/share/psi4/basis/cc-pvdz.gbs
12 atoms 2, 13, 21, 34          entry N          line 160 file
    /Users/tatianefx/psi4conda/share/psi4/basis/cc-pvdz.gbs
13 atoms 3-5, 14-15, 22-24, 35-39 entry C          line 130 file
    /Users/tatianefx/psi4conda/share/psi4/basis/cc-pvdz.gbs
14 atoms 6-11, 16-18, 25-29, 40-47 entry H          line 20 file
    /Users/tatianefx/psi4conda/share/psi4/basis/cc-pvdz.gbs
15
16 There are an even number of electrons - assuming singlet.
17 Specify the multiplicity in the molecule input block.
18
19
20 -----
21                               SCF
22 by Justin Turney, Rob Parrish, and Andy Simmonett

```



```
23                                RHF Reference
24                                1 Threads,      476 MiB Core
25                                -----
26
27 ==> Geometry <==
28
29 Molecular point group: c1
30 Full point group: C1
31
32 Geometry (in Angstrom), charge = 0, multiplicity = 1:
33
34      Center              X              Y              Z
35      Mass
36      -----
37
38      O              0.274287675591      0.995303931038
39      0.167074482057      15.994914619560
40      N              -2.385412324409      1.029503931038
41      0.753874482057      14.003074004780
42      C              -1.683412324409      -0.234796068962
43      0.965474482057      12.000000000000
44      C              -2.309412324409      -1.321196068962
45      0.101574482057      12.000000000000
46      C              -0.227912324409      -0.030896068962
47      0.602574482057      12.000000000000
48      H              -1.741512324409      -0.494896068962
49      2.027574482057      1.007825032070
50      H              -1.832912324409      -2.290696068962
51      0.283974482057      1.007825032070
52      H              -3.377212324409      -1.433896068962
53      0.319674482057      1.007825032070
54      H              -2.199712324409      -1.100896068962
55      -0.966625517943      1.007825032070
56      H              -3.366512324409      0.920803931038
57      1.009374482057      1.007825032070
58      H              -2.382012324409      1.266303931038
59      -0.238125517943      1.007825032070
60      O              -0.015812324409      -1.502496068962
61      0.456174482057      15.994914619560
62      N              0.268287675591      -0.707296068962
63      2.994174482057      14.003074004780
64      C              -0.122212324409      -2.093096068962
65      2.790674482057      12.000000000000
66      C              -0.228812324409      -2.350896068962
67      1.313074482057      12.000000000000
68      H              -1.091712324409      -2.276996068962
69      3.261574482057      1.007825032070
```

52	H	0.631087675591	-2.758696068962
		3.220574482057	1.007825032070
53	H	1.156387675591	-0.523596068962
		2.529174482057	1.007825032070
54	O	-5.589712324409	0.017203931038
		-0.827325517943	15.994914619560
55	O	-3.334912324409	-0.160596068962
		-1.043425517943	15.994914619560
56	N	-3.133212324409	2.430603931038
		-0.039225517943	14.003074004780
57	C	-4.364712324409	2.036303931038
		-0.721225517943	12.000000000000
58	C	-4.442712324409	2.719503931038
		-2.083625517943	12.000000000000
59	C	-4.331812324409	0.528403931038
		-0.880925517943	12.000000000000
60	H	-5.221812324409	2.307903931038
		-0.094725517943	1.007825032070
61	H	-4.436112324409	3.810603931038
		-1.988225517943	1.007825032070
62	H	-3.609412324409	2.425703931038
		-2.731325517943	1.007825032070
63	H	-3.055712324409	1.939703931038
		0.850574482057	1.007825032070
64	H	-5.592612324409	-0.958196068962
		-0.932425517943	1.007825032070
65	O	5.664887675591	-0.169696068962
		0.181674482057	15.994914619560
66	O	1.038987675591	-2.434696068962
		-1.618625517943	15.994914619560
67	O	5.436687675591	1.198803931038
		-1.611325517943	15.994914619560
68	O	0.932287675591	-1.971896068962
		0.603274482057	15.994914619560
69	N	3.239087675591	2.416303931038
		-0.588225517943	14.003074004780
70	C	2.761087675591	0.193503931038
		0.426474482057	12.000000000000
71	C	3.792087675591	1.309803931038
		0.189974482057	12.000000000000
72	C	2.300387675591	-0.575196068962
		-0.814925517943	12.000000000000
73	C	5.031287675591	0.802603931038
		-0.526125517943	12.000000000000
74	C	1.364587675591	-1.718896068962
		-0.513825517943	12.000000000000
75	H	1.882287675591	0.616903931038

```

      0.931574482057      1.007825032070
76      H      3.186887675591      -0.526196068962
      1.138874482057      1.007825032070
77      H      4.114687675591      1.712403931038
      1.157574482057      1.007825032070
78      H      1.783587675591      0.105603931038
      -1.500225517943      1.007825032070
79      H      3.168787675591      -0.998596068962
      -1.331625517943      1.007825032070
80      H      2.410987675591      2.782803931038
      -0.120525517943      1.007825032070
81      H      6.463787675591      -0.498096068962
      -0.283325517943      1.007825032070
82      H      0.425587675591      -3.171396068962
      -1.410225517943      1.007825032070
83
84      Running in c1 symmetry.
85
86      Rotational constants: A =      0.01236   B =      0.00406   C =
      0.00357 [cm-1]
87      Rotational constants: A =      370.60871   B =      121.59453   C =
      107.03434 [MHz]
88      Nuclear repulsion = 2542.091749210775561
89
90      Charge      = 0
91      Multiplicity = 1
92      Electrons   = 192
93      Nalpha      = 96
94      Nbeta       = 96
95
96      ==> Algorithm <==
97
98      SCF Algorithm Type is DF.
99      DIIS enabled.
100     MOM disabled.
101     Fractional occupation disabled.
102     Guess Type is SAD.
103     Energy threshold   = 1.00e-06
104     Density threshold  = 1.00e-06
105     Integral threshold = 0.00e+00
106
107     ==> Primary Basis <==
108
109     Basis Set: CC-PVDZ
110     Blend: CC-PVDZ
111     Number of shells: 216
112     Number of basis function: 460

```

```

113   Number of Cartesian functions: 485
114   Spherical Harmonics?: true
115   Max angular momentum: 2
116
117   => Loading Basis Set <=
118
119   Name: (CC-PVDZ AUX)
120   Role: JKFIT
121   Keyword: DF_BASIS_SCF
122   atoms 1, 12, 19-20, 30-33      entry 0          line 220 file
      /Users/tatianefx/psi4conda/share/psi4/basis/cc-pvdz-jkfit.gbs
123   atoms 2, 13, 21, 34          entry N          line 170 file
      /Users/tatianefx/psi4conda/share/psi4/basis/cc-pvdz-jkfit.gbs
124   atoms 3-5, 14-15, 22-24, 35-39 entry C          line 120 file
      /Users/tatianefx/psi4conda/share/psi4/basis/cc-pvdz-jkfit.gbs
125   atoms 6-11, 16-18, 25-29, 40-47 entry H          line 50 file
      /Users/tatianefx/psi4conda/share/psi4/basis/cc-pvdz-jkfit.gbs
126
127   ==> Pre-Iterations <==
128
129   -----
130   Irrep   Nso      Nmo      Nalpha   Nbeta    Ndocc   Nsocc
131   -----
132   A        460      460        0        0        0        0
133   -----
134   Total    460      460      96       96      96       0
135   -----
136
137   ==> Integral Setup <==
138
139   ==> DFJK: Density-Fitted J/K Matrices <==
140
141   J tasked:                Yes
142   K tasked:                Yes
143   wK tasked:               No
144   OpenMP threads:          1
145   Integrals threads:        1
146   Memory (MB):             357
147   Algorithm:               Disk
148   Integral Cache:          NONE
149   Schwarz Cutoff:          1E-12
150   Fitting Condition:       1E-12
151
152   => Auxiliary Basis Set <=
153
154   Basis Set: (CC-PVDZ AUX)
155   Blend: CC-PVDZ-JKFIT

```

```

156      Number of shells: 798
157      Number of basis function: 2256
158      Number of Cartesian functions: 2575
159      Spherical Harmonics?: true
160      Max angular momentum: 3
161
162      Minimum eigenvalue in the overlap matrix is 1.4362535599E-03.
163      Using Symmetric Orthogonalization.
164
165      SCF Guess: Superposition of Atomic Densities via on-the-fly atomic
        UHF.
166
167      ==> Iterations <==
168
169                                     Total Energy          Delta E          RMS |[F,P
                                     ]|
170
171      @DF-RHF iter   0: -1333.78170959988620      -1.33378e+03      2.21445e
        -02
172      @DF-RHF iter   1: -1317.64488842040009       1.61368e+01      4.07826e
        -03
173      @DF-RHF iter   2: -1283.70457811250799       3.39403e+01      1.47012e
        -02 DIIS
174      @DF-RHF iter   3: -1319.36085841447675      -3.56563e+01      1.83155e
        -03 DIIS
175      @DF-RHF iter   4: -1318.87319523789824       4.87663e-01      2.54663e
        -03 DIIS
176      @DF-RHF iter   5: -1319.03443789973062      -1.61243e-01      2.25214e
        -03 DIIS
177      @DF-RHF iter   6: -1319.97453766082481      -9.40100e-01      9.02436e
        -04 DIIS
178      @DF-RHF iter   7: -1320.06818939339996      -9.36517e-02      2.81218e
        -04 DIIS
179      @DF-RHF iter   8: -1320.08215912150490      -1.39697e-02      2.89657e
        -04 DIIS
180      @DF-RHF iter   9: -1320.09001636903008      -7.85725e-03      1.30093e
        -04 DIIS
181      @DF-RHF iter  10: -1320.09307865010965      -3.06228e-03      1.01471e
        -04 DIIS
182      @DF-RHF iter  11: -1320.09422252943614      -1.14388e-03      7.21758e
        -05 DIIS
183      @DF-RHF iter  12: -1320.09497217895023      -7.49650e-04      4.96405e
        -05 DIIS
184      @DF-RHF iter  13: -1320.09539896204615      -4.26783e-04      2.74450e
        -05 DIIS
185      @DF-RHF iter  14: -1320.09558812059640      -1.89159e-04      6.90670e
        -06 DIIS

```

```

186   @DF-RHF iter 15: -1320.09560780714946   -1.96866e-05   4.26508e
      -06 DIIS
187   @DF-RHF iter 16: -1320.09561445509758   -6.64795e-06   3.22618e
      -06 DIIS
188   @DF-RHF iter 17: -1320.09562044510130   -5.99000e-06   1.96979e
      -06 DIIS
189   @DF-RHF iter 18: -1320.09562333424628   -2.88914e-06   1.01554e
      -06 DIIS
190   @DF-RHF iter 19: -1320.09562409848650   -7.64240e-07   7.24414e
      -07 DIIS
191
192   ==> Post-Iterations <==
193
194   Orbital Energies (a.u.)
195   -----
196
197   Doubly Occupied:
198
199       1A      -20.889287       2A      -20.870954       3A      -20.729351
200       4A      -20.652285       5A      -20.644339       6A      -20.599621
201       7A      -20.585348       8A      -20.526181       9A      -15.630442
202      10A      -15.620256      11A      -15.493024      12A      -15.492759
203      13A      -11.517085      14A      -11.487661      15A      -11.429579
204      16A      -11.387198      17A      -11.365826      18A      -11.336185
205      19A      -11.307663      20A      -11.305384      21A      -11.288597
206      22A      -11.258027      23A      -11.248222      24A      -11.243974
207      25A      -11.209086      26A      -2.176731      27A      -1.568318
208      28A      -1.508638      29A      -1.497666      30A      -1.475138
209      31A      -1.436537      32A      -1.424760      33A      -1.388891
210      34A      -1.328421      35A      -1.322686      36A      -1.209720
211      37A      -1.181040      38A      -1.143404      39A      -1.138045
212      40A      -1.098620      41A      -1.087622      42A      -1.050945
213      43A      -1.012296      44A      -0.968279      45A      -0.936841
214      46A      -0.915659      47A      -0.906737      48A      -0.896635
215      49A      -0.867321      50A      -0.859251      51A      -0.831340
216      52A      -0.809712      53A      -0.799415      54A      -0.770046
217      55A      -0.764887      56A      -0.755307      57A      -0.739295
218      58A      -0.723263      59A      -0.708363      60A      -0.694887
219      61A      -0.688663      62A      -0.677096      63A      -0.676773
220      64A      -0.673325      65A      -0.660123      66A      -0.654964
221      67A      -0.652288      68A      -0.640090      69A      -0.629845
222      70A      -0.622521      71A      -0.616895      72A      -0.607051
223      73A      -0.596707      74A      -0.587945      75A      -0.579269
224      76A      -0.561443      77A      -0.551029      78A      -0.543757
225      79A      -0.540159      80A      -0.531239      81A      -0.530186
226      82A      -0.524190      83A      -0.509265      84A      -0.495298
227      85A      -0.489515      86A      -0.467818      87A      -0.460803

```

228	88A	-0.449830	89A	-0.436888	90A	-0.421249
229	91A	-0.416544	92A	-0.365458	93A	-0.362105
230	94A	-0.297810	95A	-0.249788	96A	-0.204465
231						
232	Virtual:					
233						
234	97A	-0.060623	98A	0.006946	99A	0.066458
235	100A	0.100386	101A	0.109059	102A	0.126229
236	103A	0.140293	104A	0.155220	105A	0.164055
237	106A	0.174026	107A	0.184099	108A	0.192612
238	109A	0.206768	110A	0.208847	111A	0.221153
239	112A	0.224525	113A	0.237340	114A	0.240494
240	115A	0.250954	116A	0.258559	117A	0.267281
241	118A	0.275659	119A	0.284971	120A	0.291119
242	121A	0.294076	122A	0.298767	123A	0.304191
243	124A	0.317713	125A	0.325187	126A	0.332545
244	127A	0.346225	128A	0.356964	129A	0.359953
245	130A	0.365923	131A	0.379737	132A	0.380832
246	133A	0.390430	134A	0.404497	135A	0.422093
247	136A	0.432066	137A	0.437893	138A	0.438771
248	139A	0.462421	140A	0.474779	141A	0.486117
249	142A	0.496151	143A	0.501932	144A	0.517409
250	145A	0.532302	146A	0.535213	147A	0.548000
251	148A	0.558728	149A	0.581462	150A	0.596065
252	151A	0.600879	152A	0.615656	153A	0.635388
253	154A	0.648511	155A	0.655719	156A	0.665240
254	157A	0.668751	158A	0.673467	159A	0.689058
255	160A	0.694877	161A	0.701849	162A	0.703700
256	163A	0.714918	164A	0.718374	165A	0.728668
257	166A	0.732765	167A	0.737161	168A	0.747178
258	169A	0.759583	170A	0.766564	171A	0.767597
259	172A	0.780153	173A	0.780870	174A	0.787811
260	175A	0.791845	176A	0.794157	177A	0.802886
261	178A	0.808655	179A	0.816705	180A	0.821344
262	181A	0.826160	182A	0.831130	183A	0.843524
263	184A	0.849630	185A	0.851844	186A	0.863658
264	187A	0.868425	188A	0.871826	189A	0.877441
265	190A	0.881042	191A	0.884215	192A	0.889373
266	193A	0.894999	194A	0.900662	195A	0.902261
267	196A	0.907447	197A	0.911946	198A	0.921018
268	199A	0.924423	200A	0.938628	201A	0.944233
269	202A	0.950257	203A	0.952373	204A	0.969993
270	205A	0.971429	206A	0.990242	207A	0.998148
271	208A	1.003551	209A	1.028403	210A	1.039541
272	211A	1.042464	212A	1.058939	213A	1.069746
273	214A	1.084066	215A	1.088608	216A	1.102513
274	217A	1.105467	218A	1.110336	219A	1.115227

275	220A	1.124940	221A	1.133738	222A	1.150891
276	223A	1.155683	224A	1.161340	225A	1.170932
277	226A	1.175139	227A	1.181833	228A	1.194923
278	229A	1.204578	230A	1.212711	231A	1.214982
279	232A	1.229799	233A	1.233000	234A	1.241900
280	235A	1.257279	236A	1.259465	237A	1.267560
281	238A	1.274358	239A	1.281362	240A	1.289094
282	241A	1.294924	242A	1.307271	243A	1.323576
283	244A	1.337041	245A	1.343166	246A	1.350767
284	247A	1.365089	248A	1.373311	249A	1.379651
285	250A	1.400896	251A	1.407133	252A	1.411038
286	253A	1.425350	254A	1.430060	255A	1.438098
287	256A	1.442697	257A	1.447987	258A	1.455727
288	259A	1.479356	260A	1.484923	261A	1.493926
289	262A	1.511834	263A	1.523191	264A	1.527734
290	265A	1.536180	266A	1.557280	267A	1.566947
291	268A	1.575676	269A	1.590447	270A	1.593478
292	271A	1.599591	272A	1.605298	273A	1.611056
293	274A	1.624511	275A	1.636304	276A	1.639180
294	277A	1.645939	278A	1.654205	279A	1.673934
295	280A	1.685834	281A	1.692294	282A	1.701953
296	283A	1.721198	284A	1.726767	285A	1.742017
297	286A	1.751987	287A	1.757617	288A	1.767740
298	289A	1.776955	290A	1.794203	291A	1.802432
299	292A	1.810798	293A	1.819943	294A	1.826199
300	295A	1.830412	296A	1.835301	297A	1.842822
301	298A	1.845406	299A	1.858138	300A	1.860002
302	301A	1.883289	302A	1.884629	303A	1.889511
303	304A	1.896250	305A	1.907156	306A	1.920479
304	307A	1.923102	308A	1.928161	309A	1.935809
305	310A	1.939939	311A	1.951521	312A	1.959070
306	313A	1.971490	314A	1.973267	315A	1.983304
307	316A	1.987305	317A	1.992242	318A	2.001097
308	319A	2.001914	320A	2.011870	321A	2.016332
309	322A	2.018881	323A	2.024666	324A	2.039702
310	325A	2.047510	326A	2.054500	327A	2.059547
311	328A	2.061344	329A	2.080053	330A	2.081053
312	331A	2.091100	332A	2.094245	333A	2.103579
313	334A	2.113127	335A	2.120863	336A	2.135612
314	337A	2.139959	338A	2.149846	339A	2.151265
315	340A	2.169817	341A	2.186649	342A	2.200196
316	343A	2.209283	344A	2.216745	345A	2.231006
317	346A	2.236017	347A	2.253008	348A	2.259138
318	349A	2.266845	350A	2.276379	351A	2.279761
319	352A	2.300608	353A	2.305677	354A	2.312536
320	355A	2.330275	356A	2.341157	357A	2.350038
321	358A	2.354914	359A	2.370839	360A	2.384780

322	361A	2.397674	362A	2.409610	363A	2.420307
323	364A	2.427844	365A	2.441844	366A	2.447282
324	367A	2.470384	368A	2.477641	369A	2.480440
325	370A	2.489786	371A	2.503652	372A	2.532417
326	373A	2.546238	374A	2.551563	375A	2.559205
327	376A	2.565374	377A	2.577609	378A	2.586994
328	379A	2.596908	380A	2.602752	381A	2.617035
329	382A	2.628251	383A	2.643278	384A	2.658649
330	385A	2.666474	386A	2.680028	387A	2.686577
331	388A	2.691765	389A	2.708611	390A	2.722698
332	391A	2.731686	392A	2.757730	393A	2.762307
333	394A	2.771267	395A	2.783622	396A	2.800378
334	397A	2.801997	398A	2.817636	399A	2.837475
335	400A	2.842960	401A	2.854461	402A	2.867389
336	403A	2.878199	404A	2.887639	405A	2.900460
337	406A	2.928076	407A	2.939950	408A	2.951114
338	409A	2.960562	410A	2.980920	411A	3.001066
339	412A	3.008576	413A	3.010936	414A	3.037227
340	415A	3.043174	416A	3.059947	417A	3.082868
341	418A	3.094042	419A	3.123341	420A	3.125590
342	421A	3.131399	422A	3.148782	423A	3.192612
343	424A	3.226474	425A	3.229874	426A	3.235066
344	427A	3.247602	428A	3.287497	429A	3.297844
345	430A	3.312552	431A	3.341624	432A	3.354757
346	433A	3.371891	434A	3.397810	435A	3.420641
347	436A	3.432834	437A	3.449448	438A	3.465697
348	439A	3.474243	440A	3.523583	441A	3.533080
349	442A	3.557602	443A	3.573176	444A	3.582800
350	445A	3.598443	446A	3.608253	447A	3.666238
351	448A	3.715515	449A	3.797974	450A	3.824942
352	451A	3.843372	452A	3.850326	453A	3.914544
353	454A	3.947828	455A	3.985284	456A	4.030516
354	457A	4.069648	458A	4.103402	459A	4.118220
355	460A	4.395972				
356						
357	Final Occupation by Irrep:					
358	A					
359	DOCC [96]					
360						
361	Energy converged.					
362						
363	@DF-RHF Final Energy: -1320.09562409848650					
364						
365	=> Energetics <=					
366						
367	Nuclear Repulsion Energy =			2542.0917492107755606		
368	One-Electron Energy =			-6860.2002168259050450		

```

369      Two-Electron Energy =                2998.0128435166429881
370      DFT Exchange-Correlation Energy =      0.000000000000000000
371      Empirical Dispersion Energy =          0.000000000000000000
372      PCM Polarization Energy =              0.000000000000000000
373      EFP Energy =                          0.000000000000000000
374      Total Energy =                        -1320.0956240984864962
375
376
377
378 Properties will be evaluated at    0.000000,    0.000000,    0.000000
      Bohr
379
380 Properties computed using the SCF density matrix
381
382      Nuclear Dipole Moment: (a.u.)
383      X:    -11.8631      Y:      0.7687      Z:      4.7854
384
385      Electronic Dipole Moment: (a.u.)
386      X:      9.7602      Y:     -5.9257      Z:     -3.8827
387
388      Dipole Moment: (a.u.)
389      X:     -2.1028      Y:     -5.1570      Z:      0.9027      Total:
      5.6420
390
391      Dipole Moment: (Debye)
392      X:     -5.3448      Y:    -13.1079      Z:      2.2944      Total:
      14.3405
393
394
395 *** tstop() called on Tatianes-MacBook-Air.local at Thu Jun 20
      21:39:13 2019
396 Module time:
397      user time    =      386.87 seconds =      6.45 minutes
398      system time =      14.56 seconds =      0.24 minutes
399      total time  =       247 seconds =      4.12 minutes
400 Total time:
401      user time    =    18523.43 seconds =    308.72 minutes
402      system time =     754.50 seconds =     12.57 minutes
403      total time  =    22580 seconds =    376.33 minutes

```

Listagem C.9 – timer.dat

```

1 Host: Tatianes-MacBook-Air.local
2
3 Timers On : Thu Jun 20 15:22:45 2019
4 Timers Off: Thu Jun 20 22:29:18 2019
5
6 Wall Time:    25593.00 seconds

```

```

7
8 JK: (A|Q)^-1:      1912.50u      149.65s      1202.59w      40 calls
9 JK: (A|mn) :      1497.62u      5.15s      1312.92w      363 calls
10 JK: (Q|mn) :      3772.47u      15.62s      669.36w      363 calls
11 JK: (Q|mn) Write:      4.42u      97.98s      1853.83w      203 calls
12 HF: Form H :      0.42u      0.68s      1.342858w      40 calls
13 HF: Form S/X:      14.03u      0.58s      7.730587w      40 calls
14 HF: Guess :      11.12u      0.67s      6.374783w      40 calls
15 SAD Guess :      10.82u      0.33s      5.943290w      40 calls
16 JK: (A|Q)^-1/2:      2.43u      0.02s      1.465496w      160 calls
17 JK: D :      3.22u      0.40s      1.249754w      1965 calls
18 JK: US02A0 :      0.48u      0.08s      0.355863w      1965 calls
19 JK: JK :      25159.38u      1007.32s      19849.36w      1965 calls
20 JK: J :      591.27u      3.42s      214.49w      8527 calls
21 JK: J1 :      277.15u      1.62s      112.15w      9407 calls
22 JK: J2 :      302.73u      1.57s      95.13w      9407 calls
23 JK: K :      24382.20u      116.67s      18091.65w      8527 calls
24 JK: K1 :      11878.50u      76.12s      14260.19w      9327 calls
25 JK: K2 :      12501.73u      40.45s      3830.28w      9327 calls
26 JK: A02US0 :      0.03u      0.02s      0.004958w      1965 calls
27 DIISManager::add_entry:      1.55u      2.62s      2.588620w      1925
    calls
28 DIISManager::extrapolate:      8.10u      25.53s      21.42w      1885
    calls
29 DIISManager::extrapolate: bMatrix setup:      3.17u      16.13s
    12.49w      1885 calls
30 DIISManager::extrapolate: bMatrix pseudoinverse:      0.13u
    0.00s      0.069099w      1885 calls
31 DIISManager::extrapolate: form new data:      4.73u      9.38s
    8.834411w      1885 calls
32 HF: Form G :      25164.93u      1008.03s      19852.19w      1085 calls
33 JK: (Q|mn) Read:      25.32u      695.23s      1129.63w      7647 calls
34 HF: Form F :      0.93u      0.00s      0.608197w      1085 calls
35 HF: DIIS :      92.93u      29.68s      45.67w      1085 calls
36 HF: Form C :      374.27u      11.88s      193.03w      1085 calls
37 HF: Form D :      2.80u      0.07s      1.077366w      1085 calls
38
39 *****

```