

**UNIVERSIDADE FEDERAL DE SÃO CARLOS**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA**

**“APLICAÇÃO DO MAPEAMENTO VETORIAL  
PARA BIOMOLÉCULAS”**

**ANTONIO CESAR SILVA SACCO**

**São Carlos**  
**2017**

ANTONIO CESAR SILVA SACCO

**“APLICAÇÃO DO MAPEAMENTO VETORIAL PARA BIOMOLÉCULAS”**

Tese apresentada ao Programa de  
Doutorado em Biotecnologia do  
Programa de Pós-Graduação em  
Biotecnologia da Universidade  
Federal de São Carlos.

Orientadores: **Prof<sup>a</sup> Dr<sup>a</sup> Ignez Caracelli**  
**Prof. Dr. Julio Zukerman-Schpector**

São Carlos  
2017



**UNIVERSIDADE FEDERAL DE SÃO CARLOS**

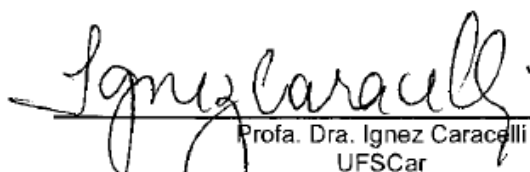
Centro de Ciências Exatas e de Tecnologia  
Programa de Pós-Graduação em Biotecnologia

---


**Folha de Aprovação**

---

Assinaturas dos membros da comissão examinadora que avaliou e aprovou a Defesa de Tese de Doutorado do candidato Antonio Cesar Silva Sacco, realizada em 09/10/2017:

  
\_\_\_\_\_  
Profa. Dra. Ignez Caracelli  
UFSCar

  
\_\_\_\_\_  
Prof. Dr. Carlos Alberto de Simone  
UFAL

  
\_\_\_\_\_  
Prof. Dr. Luis Carlos Trevelin  
UFSCar

  
\_\_\_\_\_  
Profa. Dra. Stella Hernandez Maganhi  
UFSCar

  
\_\_\_\_\_  
Profa. Dra. Regina Helena de Almeida Santos  
IQSC/USP

*Dedico este trabalho à minha amada esposa,  
Denise Maria De Lucca Ferreira Sacco.*

## **Agradecimentos**

*À minha família, Denise Maria De Lucca Ferreira Sacco, Gustavo Scherer e Patrick De Lucca, pelo incentivo e compreensão pelo tempo que esta pesquisa subtraiu da minha presença ao convívio familiar.*

*Aos meus pais, Clemente Sacco e Lélia Altina Silva pelo apoio em mais um desafio na minha vida, não furtando especial agradecimento ao meu querido irmão Aloisio César Silva Sacco pelos momentos de descontração proporcionados.*

*Aos meus orientadores, Profa. Dra. Ignez Caracelli e Prof. Dr. Julio Zukerman Schpector, pensadores que além da humildade e sapiência e que tronam uma verdadeira honra ser integrante de sua equipe, também são, juntamente comigo e com Lucas Morais de Souza, são coautores no desenvolvimento do software WIM.*

*A todos os professores do Programa de Pós-Graduação em Biotecnologia da Universidade Federal de São Carlos que contribuíram para minha formação.*

*Aos amigos do Grupo de Física Teórica de Materiais e Biomoléculas (BioMat) e do Laboratório de Cristalografia Estereodinâmica e Modelagem Molecular (LaCrEMM) da UFSCar pelo compartilhamento de experiências e conhecimento, com especial deferência a Rui Filipe Nogueira da Silva pelas críticas e sugestões referentes ao algoritmo do software para o mapeamento vetorial.*

*A todos meus sinceros agradecimentos e minha elevada estima e consideração.*

Sacco, Antonio César Silva. **Aplicação do mapeamento vetorial para biomoléculas**. 2017. Dissertação (Programa de Pós-Graduação em Biotecnologia) – UFSCar, São Carlos, 2017.

## RESUMO

O foco principal deste trabalho foi o de estudar dois tipos de interações não covalentes: as interações  $\pi$  e as ligações de hidrogênio. Este tipo de estudo é necessário quando se quer analisar, por exemplo, complexos enzima-ligante ou DNA-ligante. Com o auxílio dos programas de visualização gráfica o pesquisador termina analisando, em geral, apenas uma parte da molécula biológica em questão. Um estudo mais completo torna-se mais difícil dada a quantidade de átomos e interações existentes em moléculas biológicas. Assim, foi desenvolvido um algoritmo que permite avaliar todas as interações não covalentes  $\pi$  e ligações de hidrogênio existentes sejam intra ou intermoleculares, obtendo todas as distâncias, ângulos e átomos envolvidos, utilizando parâmetros pré-definidos ou parâmetros definidos pelo pesquisador. As interações assim obtidas podem ser visualizadas com um programa gráfico externo. Assim, foi desenvolvida a ferramenta computacional denominada WIM (Weak Interaction Mapping) que transformou o mapeamento vetorial em ferramenta de pesquisa. A automatização do processo por meio do software WIM para a pesquisa das interações não-covalentes permitiu estudar uma série de estruturas de DNA com e sem ligantes e verificar a existência de determinados padrões gerados por este tipo de interações, bem como estudar detalhadamente as deformações introduzidas pelos ligantes em diferentes estruturas de DNA. O software WIM recebeu parecer favorável no processo de proteção ao programa, Processo AIn PC 2017/004 –UFSCar, em 28/06/17.

**Palavras-chave:** mapeamento vetorial, interações  $\pi$ , interações eletrostáticas, ligações de hidrogênio.

Sacco, Antonio César Silva. **Vector mapping applied to biomolecules**. 2017. Thesis (Postgraduate Program in Biotechnology) – UFSCar, São Carlos, 2017.

## **ABSTRACT**

The main focus of this work was to study two types of non-covalent interactions:  $\pi$  interactions and hydrogen bonds. This type of study is necessary when one wants to analyze, for example, enzyme-ligand or DNA-ligand complexes. With the help of graphic visualization programs the researcher ends up analyzing, in general, only part of the studied biological molecule. A more complete study becomes more difficult given the amount of atoms and interactions in biological molecules. Thus, an algorithm was developed to evaluate all non-covalent  $\pi$  interactions and hydrogen bonds present, being intra or intermolecular, thus obtaining all the distances, angles and atoms involved, using predefined parameters or parameters defined by the researcher. The interactions so obtained can be viewed with an external graphic program. Thus, the computational tool called WIM (Weak Interaction Mapping) was developed, transforming the vector mapping into a research tool. The automation of the process using WIM software to investigate the non-covalent interactions made possible to study a series of DNA structures with and without ligands, and to verify the existence of certain patterns generated by this type of interactions, as well as to study in detail the introduced deformations by the ligands in different DNA structures. The WIM software received a favorable opinion in the process of protection of the program, Process AIn PC 2017/004 -UFSCar, on 06/27/2017.

**Keywords:** vector mapping,  $\pi$  interaction, electrostatic interactions, exploration and analysis of  $\pi$  interactions.

## LISTA DE FIGURAS

<b>Capítulo 1 - Introdução</b>		<b>página</b>
<b>Figura 1.1</b>	Representação esquemática dos parâmetros geométricos envolvidos em uma interação $\pi$ adaptado de Caracelli, Zukerman-Schpector e Tiekink, 2012.	21
<b>Figura 1.2</b>	Representação esquemática de uma tirosina e sua interação $\pi$ , com o dipolo da nuvem eletrônica sendo induzido em função de uma carga, elaborado pelo autor utilizando visualizador gráfico Discovery Studio 3.1 <sup>TM</sup>	22
<b>Figura 1.3</b>	Representação gráfica de uma interação N-H $\cdots\pi$ entre duas bases de uma molécula de DNA imagem gerada a partir de dados cristalográficos por (Malinina et al., 2002), utilizando o software Discovery Studio 3.1 <sup>TM</sup>	23
<b>Figura 1.4</b>	Representação gráfica de uma interação O-H $\cdots\pi$ em uma cyclooxygenase-2, imagem gerada a partir de dados cristalográficos (Kurumbail <i>et al.</i> , 1996) com a utilização do software Discovery Studio 3.1 <sup>TM</sup>	24
<b>Figura 1.5</b>	Interação C-H $\cdots\pi$ entre duas bases de uma molécula de DNA, imagem gerada a partir de dados cristalográficos (Drew et al., 1981), com a utilização do software Discovery Studio 3.1 <sup>TM</sup>	24
<b>Figura 1.6</b>	Nuvem eletrônica de um anel aromático que atua como doadora de elétrons para uma interação cátion $\cdots\pi$	25
<b>Figura 1.7</b>	Esquema de uma interação cátion $\cdots\pi$ , representando-se um cátion genérico posicionado sobre a face de um anel de benzeno	26
<b>Figura 1.8</b>	Representação gráfica de uma interação N $\cdots\pi$ entre as bases G 12 e G 14 de uma molécula de DNA, imagem gerada a partir de dados cristalográficos (Drew et al., 1981) e com a utilização do software Discovery Studio 3.1 <sup>TM</sup>	26
<b>Figura 1.9</b>	Representação gráfica de uma interação O $\cdots\pi$ entre a HIS219 e a PHE248 de glutathione reductase, imagem gerada a partir de dados cristalográficos (Karplus e Schulz, 1989) e visualizada com a utilização do software Discovery Studio 3.1 <sup>TM</sup>	27
<b>Figura 1.10</b>	Representações gráficas dos tipos de interações $\pi\cdots\pi$	27
<b>Figura 1.11</b>	Parâmetros geométricos para as interações $\pi\cdots\pi$ .	28
<b>Figura 1.12</b>	Representação gráfica de uma interação $\pi\cdots\pi$ entre as bases DA 18 e DT 19 de uma molécula com a imagem gerada a partir de dados cristalográficos (Drew et al., 1981) e com a utilização do software Discovery Studio 3.1 <sup>TM</sup>	28
<b>Figura 1.13</b>	Representação esquemática das ligações de hidrogênio nas estruturas secundária e terciária das proteínas	30
<b>Figura 1.14</b>	Representação esquemática das Watson-Crick <i>hydrogen bonding</i> nas ligações entre fitas do DNA efetuada pelos pares de bases G-C e A-T	30
<b>Figura 1.15</b>	Representação esquemática de uma ligação de hidrogênio, considerando como doador o átomo ligado ao hidrogênio e acceptor um segundo átomo ou centroide	31
<b>Figura 1.16</b>	Representação gráfica uma enzima	33
<b>Figura 1.17</b>	Localização dos átomos envolvidos no cálculo dos centroides de um triptofano	36
<b>Figura 1.18</b>	Representação gráfica dos vetores com origem no centroide de um anel com destino em átomos localizados à uma distância máxima de 4,5 Å	37
<b>Figura 1.19</b>	Determinação do vetor normal ao plano do anel aromático de resíduo de	38



	aminoácido tirosina	
<b>Figura 1.20</b>	Determinação do ângulo $\alpha$ entre o vetor normal $\vec{n}$ ao plano $\pi$ formado pelo anel aromático e o vetor distância $\vec{v}$ com base na estrutura de uma tirosina	39
<b>Figura 1.21</b>	Fluxograma das operações envolvidas no mapeamento vetorial para as interações $\pi$	40
<b>Figura 1.22</b>	Representação dos vetores e o ângulo $\phi$ entre estes, utilizados para a localização de ligações de hidrogênio utilizando o Mapeamento Vetorial	41
<b>Figura 1.23</b>	Fluxograma resumido das operações envolvidas no mapeamento vetorial para a localização das ligações de hidrogênio	42

## **Capítulo 2 – Materiais e métodos**

**página**

<b>Figura 2.1</b>	Imagem da tela da planilha eletrônica construída para o cálculo da coordenada do centroide e determinação do vetor normal ao plano do anel de uma base Citosina	45
<b>Figura 2.2</b>	Planilha com a relação das coordenadas dos centroides e vetores normais ao plano de cada anel aromático calculados para o DNA – SL	46
<b>Figura 2.3</b>	Imagem da planilha eletrônica elaborada para o cálculo do módulo do vetor distância e os ângulos entre o vetor distância e o vetor normal para cada relação centroide-átomo e centroide-centroide da estrutura em estudo	47
<b>Figura 2.4</b>	Imagem da planilha eletrônica utilizada para exportar os resultados válidos uma interação $\pi$ , na em arquivo compatível para visualização em software gráfico	47
<b>Figura 2.5</b>	Visualização da imagem da interação $N\cdots\pi$ obtida por meio de cálculo vetorial de estrutura de DNA (Drew et al., 1981)	48
<b>Figura 2.6</b>	Visualização da interação $N\cdots\pi$ obtida por meio de cálculo vetorial de estrutura de DNA	48
<b>Figura 2.7</b>	Visualização da estrutura de DNA (Drew et al., 1981) com todas as interações localizadas pelo mapeamento vetorial	49
<b>Figura 2.8</b>	Imagens das estruturas de DNA com as interações $\pi$ identificadas por mapeamento vetorial	53
<b>Figura 2.9</b>	Visualização dos centroides da estrutura DNA – SL (Drew et al., 1981) obtidos por cálculo vetorial utilizando planilha eletrônica do Microsoft Excel	54
<b>Figura 2.10</b>	Visualização gráfica da deformação relativa calculada em função da presença de ligante no sulco da estrutura DNA – LS, tendo como referência a estrutura DNA – SL	56
<b>Figura 2.11</b>	Visualização gráfica da deformação relativa calculada em função da presença de ligante intercalado na estrutura DNA – LI, tendo como referência a estrutura DNA – SL	57
<b>Figura 2.12</b>	<b>2.11</b> Diagrama de transferência de cargas eletrostáticas nos sistemas $\pi$ nas estruturas de DNA	59
<b>Figura 2.13</b>	Representação gráfica das interações $\pi$ entre o ligante e as bases de uma estrutura de DNA – LS	61
<b>Figura 2.14</b>	Representação gráfica da tabulação das distâncias entre os átomos de hidrogênio e centroides nas interações $C-H\cdots\pi$ da cadeia lateral de proteínas	62
<b>Figura 2.15</b>	Representação gráfica da tabulação das distâncias das interações $N\cdots\pi$ das	63

	estruturas DNA – SL, DNA – LS e DNA – LI	
<b>Figura 2.16</b>	Representação gráfica da tabulação dos ângulos entre o vetor distância e o vetor normal ao plano do anel aromático das interações N $\cdots$ $\pi$ das estruturas DNA – SL, DNA – LS e DNA – LI	64
<b>Figura 2.17</b>	Tela inicial do WIM	69
<b>Figura 2.18</b>	Tela de criação de configuração para o cálculo das interações $\pi$	70
<b>Figura 2.19</b>	Tela de configuração ajustada para os parâmetros personalizados	70
<b>Figura 2.20</b>	Ilustração da visualização das interações $\pi$ resultantes do processamento pelo WIM	71

### **Capítulo 3 – Resultados e Discussões**

		<b>página</b>
<b>Figura 3.1</b>	Watson Crick, pares de bases, A-T e G-C	78
<b>Figura 3.2</b>	Distribuição relativa das estruturas de DNA em função das suas características	79
<b>Figura 3.3</b>	Representação gráfica de parte de uma molécula de DNA (Drew et al., 1981) com as interações $\pi$ existentes entre as bases, determinadas vetorialmente	85
<b>Figura 3.4</b>	Interações $\pi$ existentes entre as bases de fitas diferentes mapeadas em estruturas sem ligantes	91
<b>Figura 3.5</b>	4 Interações $\pi$ existentes entre as bases de fitas diferentes mapeadas em estrutura com ligante intercalado (Malinina et al., 2002)	95
<b>Figura 3.6</b>	Interações $\pi$ existentes entre as bases de fitas diferentes mapeadas em estruturas com ligantes no sulco	100
<b>Figura 3.7</b>	Imagens das estruturas de DNA com as interações $\pi$ entre fitas identificadas por mapeamento vetorial	101
<b>Figura 3.8</b>	Distribuição dos tipos de ligações entre o DNA e os seus ligantes com anéis aromáticos no sulco	106

## LISTA DE TABELAS

<b>Capítulo 1 - Introdução</b>		<b>página</b>
<b>Tabela 1.1</b>	Forças eletrostáticas combinadas	20
<b>Tabela 1.2</b>	Parâmetros geométricos teóricos para as interações $\pi$ , por tipo	29
<b>Tabela 1.3</b>	Aceptores e doadores de elétrons em ligações de hidrogênio	31
<b>Tabela 1.4</b>	Parâmetros geométricos das ligações de hidrogênio representados na Figura 1.13	32
<b>Tabela 1.5</b>	Coordenadas espaciais da Tyr 23 existente na estrutura cristalográfica de uma glutatona redutase (Karplus e Schulz, 1989)	34
 <b>Capítulo 2 – Materiais e métodos</b>		 <b>página</b>
<b>Tabela 2.1</b>	Interações $\pi$ identificadas em cada estrutura de DNA mapeada, incluídas as interações presentes nos ligantes.	52
<b>Tabela 2.2</b>	Interações $\pi$ identificadas em cada estrutura de DNA mapeada, incluídas as interações presentes nos ligantes	52
<b>Tabela 2.3</b>	Valor do deslocamento dos centroides da estrutura DNA – LS, tendo como referência a posição do mesmo centroide na estrutura DNA – SL Valor do deslocamento dos centroides da estrutura DNA – LS, tendo como referência a posição do mesmo centroide na estrutura DNA – SL	55
<b>Tabela 2.4</b>	Valor do deslocamento dos centroides da estrutura DNA – LI, tendo como referência a posição do mesmo centroide na estrutura DNA – SL	57
<b>Tabela 2.5</b>	Interações $\pi$ com origem na base G2, anel aromático com 6 átomos, com os átomos e centroides relacionados por bases com o respectivo comprimento $d$ da interação e ângulo $\alpha$	60
<b>Tabela 2.6</b>	Interações $\pi$ com origem na base T20, com os átomos e centroides relacionados por bases com o respectivo comprimento $d$ da interação e ângulo $\alpha$	61
<b>Tabela 2.7</b>	Trabalhos de pesquisa, produção acadêmica e programa de computador decorrentes do Mapeamento Vetorial	65
<b>Tabela 2.8</b>	Tempo de processamento das estruturas cristalográficas de DNA em um notebook DELL modelo Inspiron 15R SE	75
 <b>Capítulo 3 – Resultados e Discussões</b>		 <b>página</b>
<b>Tabela 3.1</b>	Estruturas de DNA sem ligantes	80
<b>Tabela 3.2</b>	Estrutura de DNA com ligante intercalado	80
<b>Tabela 3.3</b>	Estrutura de DNA com ligantes no sulco	80
<b>Tabela 3.4</b>	Interações $\pi$ entre fitas mapeadas pelo WIM para estrutura de DNA sem ligante (Drew et al., 1981)	86
<b>Tabela 3.5</b>	Interações $\pi$ entre fitas mapeadas pelo WIM para as estruturas de DNA sem ligantes, quantificação e localização das interações entre fitas, possuindo como destino átomos da Guanina 2	87
<b>Tabela 3.6</b>	Interações $\pi$ entre fitas mapeadas pelo WIM para as estruturas de DNA sem	87

	ligantes, quantificação e localização das interações entre fitas, possuindo como destino átomos e centroide da Guanina 4	
<b>Tabela 3.7</b>	Interações $\pi$ entre fitas mapeadas pelo WIM para as estruturas de DNA sem ligantes, quantificação e localização das interações entre fitas, possuindo como destino átomos e centroide da Guanina 10	87
<b>Tabela 3.8</b>	Interações $\pi$ entre fitas mapeadas pelo WIM para as estruturas de DNA sem ligantes, quantificação e localização das interações entre fitas, possuindo como destino átomos e centroide da Guanina 12	88
<b>Tabela 3.9</b>	Interações $\pi$ entre fitas mapeadas pelo WIM para as estruturas de DNA sem ligantes, quantificação e localização das interações entre fitas, possuindo como destino átomos da Guanina 14	88
<b>Tabela 3.10</b>	Interações $\pi$ entre fitas mapeadas pelo WIM para as estruturas de DNA sem ligantes, quantificação e localização das interações entre fitas, possuindo como destino átomos e centroide da Guanina 16	88
<b>Tabela 3.11</b>	Interações $\pi$ entre fitas mapeadas pelo WIM para as estruturas de DNA sem ligantes, quantificação e localização das interações entre fitas, possuindo como destino átomos e centroide da Guanina 22	89
<b>Tabela 3.12</b>	Interações $\pi$ entre fitas mapeadas pelo WIM para as estruturas de DNA sem ligantes, quantificação e localização das interações entre fitas, possuindo como destino átomos e centroide da Guanina 24	89
<b>Tabela 3.13</b>	Distribuição em valores absolutos das interações $\pi$ entre fitas mapeadas pelo WIM para as estruturas de DNA sem ligantes, evidenciando as ligações existentes entre fitas diferentes	89
<b>Tabela 3.14</b>	Distribuição em valores relativos das interações $\pi$ entre fitas mapeadas pelo WIM para as estruturas de DNA sem ligantes, evidenciando as ligações existentes entre fitas diferentes	89
<b>Tabela 3.15</b>	Interações $\pi$ entre fitas mapeadas pelo WIM para as estruturas de DNA com ligante intercalado, quantificação e localização das interações entre fitas	92
<b>Tabela 3.16</b>	Interações $\pi$ entre fitas mapeadas pelo WIM para as estruturas de DNA com ligante intercalado, quantificação e localização das interações entre fitas, possuindo como destino átomos da Guanina 14	92
<b>Tabela 3.17</b>	Interações $\pi$ entre fitas mapeadas pelo WIM para as estruturas de DNA com ligante intercalado, quantificação e localização da interação entre fitas, possuindo como destino átomo da Guanina 16	92
<b>Tabela 3.18</b>	Interações $\pi$ entre fitas mapeadas pelo WIM para as estruturas de DNA com ligante intercalado, quantificação e localização da interação entre fitas, possuindo como destino átomo da Guanina 22	93
<b>Tabela 3.19</b>	Interações $\pi$ entre fitas mapeadas pelo WIM para as estruturas de DNA com ligante intercalado, quantificação e localização das interações entre fitas, possuindo como destino átomos da Guanina 24	93
<b>Tabela 3.20</b>	Distribuição em valores absolutos das interações $\pi$ entre fitas mapeadas pelo WIM para a estrutura de DNA com ligante intercalado, evidenciando as ligações existentes entre fitas diferentes	94
<b>Tabela 3.21</b>	Distribuição em valores relativos das interações $\pi$ entre fitas mapeadas pelo WIM para a estrutura de DNA com ligante intercalado, evidenciando as ligações existentes entre fitas diferentes	94

<b>Tabela 3.22</b>	Interações $\pi$ entre fitas mapeadas pelo WIM para as estruturas de DNA com ligantes no sulco, quantificação e localização das interações entre fitas, possuindo como destino átomos da Guanina 2	96
<b>Tabela 3.23</b>	Interações $\pi$ entre fitas mapeadas pelo WIM para as estruturas de DNA com ligantes no sulco, quantificação e localização das interações entre fitas, possuindo como destino átomos e centroide da Guanina 4	96
<b>Tabela 3.24</b>	Interações $\pi$ entre fitas mapeadas pelo WIM para as estruturas de DNA com ligantes no sulco, quantificação e localização das interações entre fitas, possuindo como destino átomos e centroide da Guanina 10	97
<b>Tabela 3.25</b>	Interações $\pi$ entre fitas mapeadas pelo WIM para as estruturas de DNA com ligantes no sulco, quantificação e localização das interações entre fitas, possuindo como destino átomos e centroide da Guanina 12	97
<b>Tabela 3.26</b>	Interações $\pi$ entre fitas mapeadas pelo WIM para as estruturas de DNA com ligantes no sulco, quantificação e localização das interações entre fitas, possuindo como destino átomos da Guanina 14	97
<b>Tabela 3.27</b>	Interações $\pi$ entre fitas mapeadas pelo WIM para as estruturas de DNA com ligantes no sulco, quantificação e localização das interações entre fitas, possuindo como destino átomos e centroide da Guanina 16	98
<b>Tabela 3.28</b>	Interações $\pi$ entre fitas mapeadas pelo WIM para as estruturas de DNA com ligantes no sulco, quantificação e localização das interações entre fitas, possuindo como destino átomos e centroide da Guanina 22	98
<b>Tabela 3.29</b>	Interações $\pi$ entre fitas mapeadas pelo WIM para as estruturas de DNA com ligantes no sulco, quantificação e localização das interações entre fitas, possuindo como destino átomos e centroide da Guanina 24	98
<b>Tabela 3.30</b>	Distribuição em valores absolutos das interações $\pi$ entre fitas mapeadas pelo WIM para as estruturas de DNA com ligantes no sulco, evidenciando as ligações existentes entre fitas diferentes	99
<b>Tabela 3.31</b>	Distribuição em valores relativos das interações $\pi$ entre fitas mapeadas pelo WIM para as estruturas de DNA com ligantes no sulco, evidenciando as ligações existentes entre fitas diferentes	99
<b>Tabela 3.32</b>	Distribuição em valores absolutos das interações $\pi$ entre fitas mapeadas pelo WIM para as estruturas de DNA com sequência de nucleotídeos (*CP*GP*CP*GP*AP*AP*TP*TP*CP*GP*CP*G), evidenciando as ligações existentes entre fitas diferentes	102
<b>Tabela 3.33</b>	Distribuição em valores relativos das interações $\pi$ entre fitas mapeadas pelo WIM para as estruturas de DNA com sequência de nucleotídeos (*CP*GP*CP*GP*AP*AP*TP*TP*CP*GP*CP*G), evidenciando as ligações existentes entre fitas diferentes	102
<b>Tabela 3.34</b>	Relação existente entre as interações $\pi$ e ligações de hidrogênio e as origens destas para a estrutura com o ligante intercalado	103
<b>Tabela 3.35</b>	Estruturas de DNA com íons sequestrados por moléculas de água no sulco menor	104
<b>Tabela 3.36</b>	Estruturas de DNA com compostos à base de platina	104
<b>Tabela 3.37</b>	Estruturas de DNA com os ligantes dotados de anéis aromáticos localizados no sulco menor e interações mapeadas	105
<b>Tabela 3.38</b>	Relações das interações não covalentes mapeadas pelo WIM para a estrutura de DNA com o ligante intercalado	107

<b>Tabela 3.39</b>	Relações das interações não covalentes mapeadas pelo WIM para as estruturas de DNA sem ligantes	108
<b>Tabela 3.40</b>	Relações das interações não covalentes mapeadas pelo WIM para as estruturas de DNA íons aprisionados por cadeia de moléculas de H <sub>2</sub> O no sulco menor	108
<b>Tabela 3.41</b>	Relações das interações não covalentes mapeadas pelo WIM para as estruturas de DNA com os ligantes sem anéis aromáticos no sulco menor	108
<b>Tabela 3.42</b>	Relações das interações não covalentes mapeadas pelo WIM para as estruturas de DNA com os ligantes dotados de anéis aromáticos no sulco menor	109
<b>Tabela 3.43</b>	Relações das interações não covalentes mapeadas pelo WIM para as estruturas de DNA	110

## LISTA DE ABREVIATURAS E SIGLAS

°	<i>Graus</i>
Å	<i>Ångstrom</i>
BioMat	<i>Grupo de Física Teórica de Materiais e Biomoléculas</i>
C#	<i>Linguagem de programação</i>
CPU	<i>Central Processing Unit</i>
DNA	<i>Deoxyribonucleic acid (ácido desoxirribonucléico)</i>
HB	<i>Hydrogen Bond</i>
LaCrEMM	<i>Laboratório de Cristalografia Estereodinâmica e Modelagem Molecular</i>
LP	<i>Linguagem de programação</i>
PDB	<i>Protein Data Bank</i>
PDBSum	<i>Protein Data Bank Summary</i>
RSCB	<i>Research Collaboratory for Structure Bioinformatics</i>

### Lista de nucleotídeos:

DC	Citosina	DA	Adenina
DT	Timina	DG	Guanina

### Lista de aminoácidos utilizados:

HIS	Histidina	TRP	Triptofano
PHE	Fenilalanina	TYR	Tirosina

# SUMÁRIO

Resumo

Abstract

Lista de Figuras

Lista de Tabelas

Lista de Abreviaturas e Siglas

Lista de Anexos

## Apresentação

<b>Capítulo 1 - Introdução</b>		<b>página</b>
<b>1</b>	<b>INTERAÇÕES NÃO COVALENTES</b>	<b>19</b>
<b>1.1</b>	<b>Forças Eletrostáticas</b>	<b>20</b>
<b>1.1.1</b>	<b>Interações <math>\pi</math></b>	<b>20</b>
<b>1.1.2</b>	<b>Tipos de interações <math>\pi</math></b>	<b>22</b>
<b>1.1.3</b>	<b>Interações O-H<math>\cdots\pi</math>, N-H<math>\cdots\pi</math> e C-H<math>\cdots\pi</math></b>	<b>23</b>
<b>1.1.4</b>	<b>Interações cátion<math>\cdots\pi</math>, O<math>\cdots\pi</math> e N<math>\cdots\pi</math></b>	<b>25</b>
<b>1.1.5</b>	<b>Interações <math>\pi\cdots\pi</math></b>	<b>27</b>
<b>1.1.6</b>	<b>Parâmetros geométricos das interações <math>\pi\cdots\pi</math></b>	<b>29</b>
<b>1.2</b>	<b>Ligações de Hidrogênio</b>	<b>29</b>
<b>1.2.1</b>	<b>Parâmetros geométricos das Ligações de Hidrogênio</b>	<b>31</b>
<b>1.3</b>	<b>O Mapeamento Vetorial</b>	<b>32</b>
<b>1.3.1</b>	<b>Operações com vetores</b>	<b>35</b>
<b>1.3.2</b>	<b>O cálculo vetorial aplicado às biomoléculas</b>	<b>35</b>
<b>1.3.2.1</b>	<b>Interações <math>\pi</math></b>	<b>35</b>
<b>1.3.2.2</b>	<b>Ligações de Hidrogênio</b>	<b>35</b>
<b>Capítulo 2 – Materiais e métodos</b>		<b>página</b>
<b>2</b>	<b>A APLICAÇÃO DO MAPEAMENTO VETORIAL – O INÍCIO</b>	<b>44</b>
<b>2.1</b>	<b>Planilhas eletrônicas – Os primeiros resultados</b>	<b>50</b>
<b>2.1.1</b>	<b>Análises quantitativa e qualitativa de interações <math>\pi</math> em estruturas de DNA</b>	<b>51</b>
<b>2.1.2</b>	<b>Deformação da estrutura do DNA causada pela presença de ligantes</b>	<b>53</b>
<b>2.1.3</b>	<b>Visualização dos caminhos eletrônicos formados pelas interações <math>\pi</math></b>	<b>58</b>
<b>2.1.4</b>	<b>Variação das interações <math>\pi</math> em função da presença e tipo de ligantes</b>	<b>59</b>



2.1.5	Busca por padrões de interações $\pi$ em estruturas de DNA	62
2.2	O Mapeamento Vetorial na pesquisa e produção acadêmicas	64
2.2.1	Projetos de pesquisa	65
2.2.2	Trabalhos apresentados e publicados	66
2.2.3	Programa de computador	67
2.3	Desenvolvimento de um <i>software</i> para o Mapeamento Vetorial	67
2.3.1	Requisitos básicos do <i>software</i>	68
2.3.2	Definição da linguagem de programação	68
2.3.3	Interface com o usuário	69
2.3.4	<i>Output</i> do processamento	71
2.3.5	Considerações sobre o desenvolvimento	72
2.3.5.1	Conversão de arquivos .PDB em .MOL2	72
2.3.5.2	Identificação e localização dos anéis aromáticos	72
2.3.5.3	Localização das ligações de hidrogênio	73
2.3.5.4	Formatos numéricos	74
2.3.5.5	Tempo de processamento	75
2.3.5.6	O WIM e a inovação	76
 <b>Capítulo 3 – Resultados e Discussões</b>		<b>página</b>
3	WIM – RESULTADOS OBTIDOS EM ESTRUTURAS DE DNA	78
3.1	Definição das estruturas	79
3.2	Resultados obtidos	84
3.3	Papel das interações $\pi$ na conformação da estrutura do DNA	84
3.3.1	Estruturas de DNA sem ligantes	85
3.3.2	Estrutura de DNA com ligante intercalado	92
3.3.3	Estruturas de DNA com ligantes no sulco	96
3.3.4	Visão geral das interações $\pi$	101
3.4	Interações entre os ligantes e as estruturas do DNA	103
3.5	Relação entre interações $\pi$ e ligações de hidrogênio na estrutura de DNA	107
 <b>Capítulo 4 – Conclusões</b>		<b>página</b>
	Conclusões	113
 <b>Referências</b>		<b>página</b>
	Referências	118

# Capítulo 1

---

INTRODUÇÃO

O objetivo deste capítulo é proporcionar uma visão geral sobre a pesquisa desenvolvida. Efetuou-se uma revisão bibliográfica a respeito dos conceitos inerentes a duas interações não covalentes específicas: as interações  $\pi$ , mostrando sua tipologia e características específicas e as ligações de hidrogênio. Para ambos os casos, além da exposição das suas relevâncias para os sistemas biológicos, são mostrados os parâmetros geométricos para que tanto as interações  $\pi$  quanto as ligações de hidrogênio possam existir. Uma vez conhecendo estas duas interações não covalentes e seus respectivos parâmetros, será explanado o conceito do mapeamento vetorial criado pelo autor durante a sua pesquisa de mestrado.

## 1 INTERAÇÕES NÃO COVALENTES

Segundo Lehn (1995), a química supramolecular consiste no estudo dos arranjos moleculares e ligações intermoleculares que geram moléculas de grande complexidade, resultantes da associação de duas ou mais espécies químicas unidas por forças não-covalentes intermoleculares. Dentro deste contexto, desde 1995, o Laboratório de Cristalografia, Estereodinâmica e Modelagem Molecular (Zukerman-Schpector, J.; Caracelli, I.; Haiduc, I., 2012) do Departamento de Química da Universidade Federal de São Carlos e o BioMat - Grupo de Física Teórica de Materiais e Biomoléculas (Caracelli, Zukerman-Schpector e Araujo-Moreira, 2012) do Departamento de Física da Universidade Federal de São Carlos estudam estas interações em biomoléculas e interações receptor-fármaco, com trabalhos envolvendo proteínas (Cunha, R. L. O. R. et al., 2006; Vega-Teijido, M.; Caracelli, I.; Zukerman-Schpector, J., 2006; Caracelli, I. et al, 2010; Caracelli, Zukerman-Schpector e Tiekink, 2012) e também com estruturas de DNA (Seng et al., 2008; Seng et al., 2010; Sacco, 2013; Caracelli, I. et al., 2014; Caracelli, I. et al., 2015; Sacco, A. C. S. et al., 2016).

## 1.1 FORÇAS ELETROSTÁTICAS

As interações não covalentes são de natureza eletrostática e subdividem-se em forças eletrostáticas simples e forças eletrostáticas combinadas. As forças eletrostáticas simples são resultantes das propriedades da partícula e também de fatores geométricos envolvendo a posição das partículas no sistema (Ermondi e Caron, 2006; Sacco, 2013). Já as forças eletrostáticas combinadas envolvem as forças eletrostáticas simples e fenômenos quânticos, são exemplos destes tipos de interações as ligações de Hidrogênio e os sistemas  $\pi$  (Tabela 1.1).

**Tabela 1.1** Forças eletrostáticas combinadas, adaptado de Ermondi e Caron (2006)

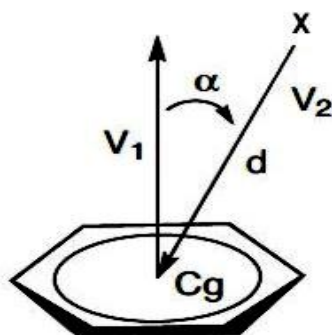
Categoria	Classificação e energia aproximada por interação	
Ligações de Hidrogênio	Entre o Hidrogênio e átomos com baixa eletronegatividade	4 a 60 kJ.mol <sup>-1</sup>
	Entre o Hidrogênio e átomos com alta eletronegatividade	
Envolvendo sistemas $\pi$	Face a face Paralela – deslocada	5 a 80 kJ.mol <sup>-1</sup>
	Em forma de “T” Cátion $\cdots \pi$	até 50 kJ.mol <sup>-1</sup>

### 1.1.1 INTERAÇÕES $\pi$

A existência de uma interação  $\pi$  está relacionada à presença de anéis aromáticos e/ou heteroaromáticos do ligante que interagem com as cadeias laterais de aminoácidos ou seus anéis e também (Meyer et al., 2003; Tiekink e Zukerman-Schpector, 2011) na manutenção da estrutura do DNA, tanto pela interação entre as suas bases como também com eventuais ligantes (Mignon et al., 2005; Ermondi e Caron, 2006; Seng et al., 2008; Tewari e Dubey, 2008). Diferentemente das ligações covalentes cujas distâncias envolvidas não ultrapassam os 2,0Å, as interações  $\pi$  apresentam requisitos geométricos mais complexos (Kiviniemi, 2001; Babu, 2003) que além de possuírem distâncias maiores, incluem o ângulo entre a interação e o plano do anel (Babu, 2003; Caracelli, Zukerman-Schpector e Tiekink, 2012), ângulo esse não superior a 30°. Ainda conforme Sacco (2013),

estas interações, por possuírem a origem no centroide do plano do anel, também são caracterizadas por sua direcionalidade. A configuração típica de uma interação  $\pi$  pode ser visualizada na Figura 1.1, onde são apresentados os parâmetros que foram utilizados na caracterização de uma interação  $\pi$ .

- C<sub>g</sub>** : centroide do anel;
- X** : átomo ou centroide envolvido na interação  $\pi$  ;
- V<sub>1</sub>** : vetor normal ao plano do anel, com origem no centroide;
- V<sub>2</sub>** : vetor que vai do centroide ao átomo sujeito à interação  $\pi$  ou um outro centroide;
- d** : módulo de V<sub>2</sub> ou a distância do centroide ao átomo sujeito à interação  $\pi$  ou um outro centroide;
- $\alpha$**  : ângulo entre V<sub>2</sub> e o vetor normal (V<sub>1</sub>).

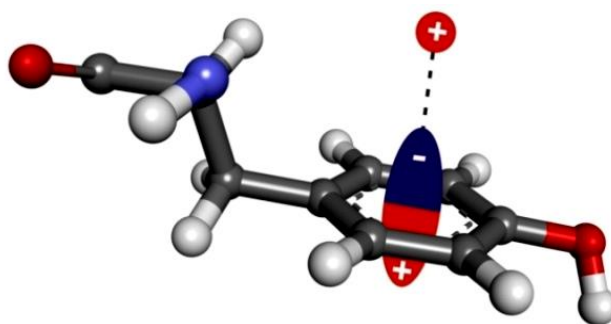


**Figura 1.1** Representação esquemática dos parâmetros geométricos envolvidos em uma interação  $\pi$ , adaptado de Caracelli, Zukerman-Schpector e Tiekink, 2012.

Segundo Sacco, Caracelli e Zukerman-Schpector (2010) e outros autores (Meyer et al., 2003; Ermondi e Caron, 2006; Shaikh e Jayaram, 2012), a compreensão das interações  $\pi$  é fundamental para a pesquisa e desenvolvimento de fármacos pela compreensão das interações entre proteína-ligante e DNA-ligante e também a sua função como estabilizadora das estruturas de biomoléculas (Mcgaughey et al., 1998; Wintjens et al., 2000; Sherrill et al., 2009; Shanthi et al., 2010).

### 1.1.2 TIPOS DE INTERAÇÕES $\pi$

Conforme exposto anteriormente, os sistemas  $\pi$ , além da posição relativa dos anéis na interação, sofrem influência dos átomos ou heteroátomos próximos (Meyer *et al.*, 2003; Ermondi e Caron, 2006). A atração eletrostática exercida por átomos carregados positiva ou negativamente ou o alinhamento de dipolos que estejam localizados em posições espaciais específicas polarizam a nuvem eletrônica do anel aromático, formando um dipolo induzido, este fenômeno pode ser observado na Figura 1.2 (Hunter e Sanders, 1990; Hunter *et al.*, 2001; Sacco, 2013).

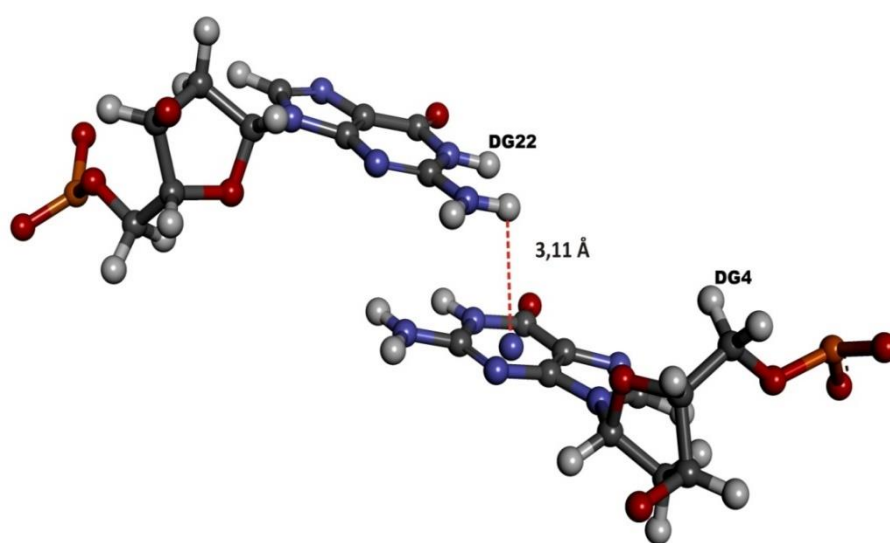


**Figura 1.2** Representação esquemática de uma tirosina e sua interação  $\pi$ , com o dipolo da nuvem eletrônica sendo induzido em função de uma carga, elaborado pelo autor utilizando visualizador gráfico Discovery Studio 3.1™ (Sacco, 2013).

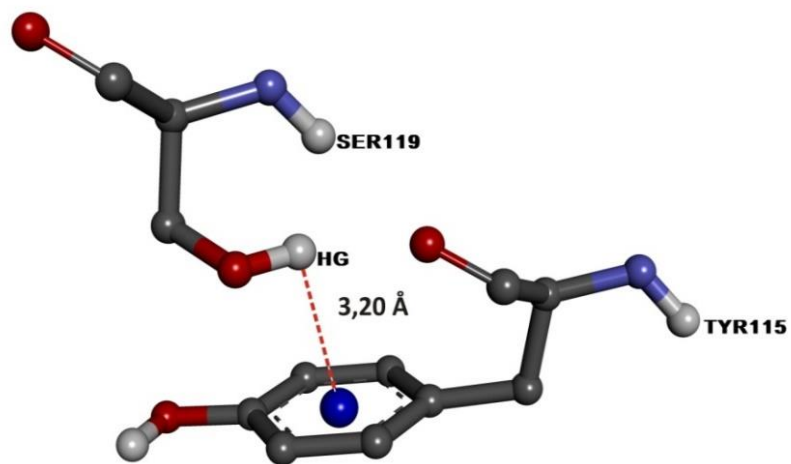
Dentro deste contexto, segundo Tewari e Dubey (2008), as interações mais comuns nos sistemas biológicos são as  $\pi \cdots \pi$ , O-H $\cdots\pi$ , N-H $\cdots\pi$  e cátion $\cdots\pi$ , sendo que incluem-se nesse grupo, conforme Imai (Imai *et al.*, 2009), as interações O $\cdots\pi$  e N $\cdots\pi$ . A seguir, de forma sucinta, serão mostradas as características dessas interações. Considerando-se que para sua existência as interações  $\pi$  sejam dependentes de anéis aromáticos, em biomoléculas estão associados aos aminoácidos Fenilalanina (Phe), Histidina (His), Triptofano (Trp) e Tyrosina (Tyr), e ainda às bases do DNA: Adenina (DA), Citosina (DC), Glutamina (DG) e Timina (DT). Podem ainda estar presentes em ligantes que possuam anéis aromáticos, (Mcgaughey *et al.*, 1998).

### 1.1.3 INTERAÇÕES O-H $\cdots$ $\pi$ , N-H $\cdots$ $\pi$ e C-H $\cdots$ $\pi$

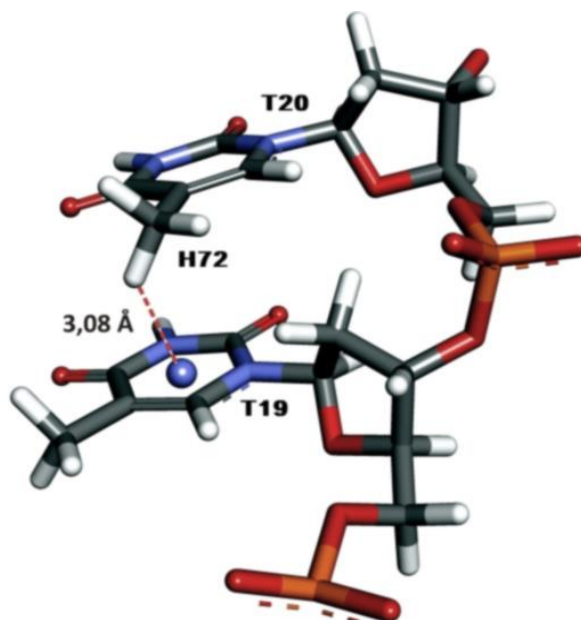
É possível tratar as interações X-H $\cdots$  $\pi$  como sendo interações onde a nuvem eletrônica do anel aromático (aceptor) interage com um átomo de hidrogênio de um grupo doador X-H (sendo X = O, N, S), de forma análoga ao que ocorre nas ligações de hidrogênio (Tsuzuki et al., 2000; Sherrill et al., 2009; Mohan et al., 2010; Nishio, 2011; Nishio et al., 2012). Todavia, conforme observado por Nishio (2011), as interações N-H $\cdots$  $\pi$  e O-H $\cdots$  $\pi$  são mais fracas do que as ligações de hidrogênio e ainda de acordo com Mohan e colaboradores (2010), essas interações, juntamente com as interações C-H $\cdots$  $\pi$ , são responsáveis pela estabilização das proteínas e para efetuar o reconhecimento biológico. Exemplos destas interações são ilustrados nas Figuras 1.3, 1.4 e 1.5.



**Figura 1.3** Representação gráfica de uma interação N-H $\cdots$  $\pi$  entre duas bases de uma molécula de DNA imagem gerada a partir de dados cristalográficos por (Malinina et al., 2002), utilizando o software Discovery Studio 3.1<sup>TM</sup>, extraído de Sacco (2013).



**Figura 1.4** Representação gráfica de uma interação O-H $\cdots$  $\pi$  em uma cyclooxygenase-2, imagem gerada a partir de dados cristalográficos (Kurumbail et al., 1996) com a utilização do software Discovery Studio 3.1<sup>TM</sup>, extraído de Sacco (2013).



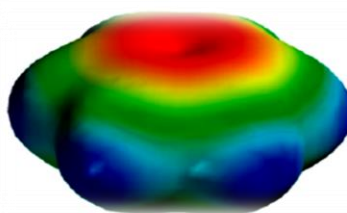
**Figura 1.5** Interação C-H $\cdots$  $\pi$  entre duas bases de uma molécula de DNA, imagem gerada a partir de dados cristalográficos (Drew et al., 1981), com a utilização do software Discovery Studio 3.1<sup>TM</sup>, extraído de Sacco (2013).



#### 1.1.4 INTERAÇÕES CÁTION $\cdots\pi$ , O $\cdots\pi$ e N $\cdots\pi$ em ANÉIS AROMÁTICOS

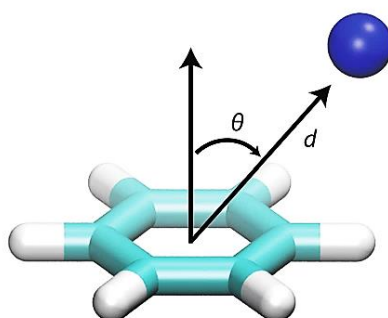
Conforme exposto por Pinheiro e coautores (Pinheiro et al., 2017), interações não covalentes do tipo Cátion $\cdots\pi$  ocorrem entre os anéis aromáticos (sistema  $\pi$ ) e cátions ou grupos positivamente carregados em sua adjacência. Segundo a literatura (Ma e Dougherty, 1997; Gallivan e Dougherty, 1999; Dougherty, 2007; Du, Long, et al., 2012; Du, Meng, et al., 2012), em meios biológicos estas interações envolvem como sistema  $\pi$  a Tirosina (Tyr), o Triptofano (Trp), a Fenilalanina (Phe) e a Histidina (His) e como cátions orgânicos os aminoácidos Fenilalanina (Phe), Tirosina (Tyr), Triptofano (Trp), Arginina (Arg), Lisina (Lys) e Histidina (His), além de todos os cátions metálicos, como por exemplo do Potássio ( $K^+$ ), Sódio ( $Na^+$ ) e Cálcio ( $Ca^{2+}$ ). Além dos cátions citados, em trabalhos de grande repercussão, Zukerman-Schpector (Zukerman-Schpector e Haiduc, 2002; Tiekink e Zukerman-Schpector, 2009a; b) identificou esse tipo de interação em estruturas que contém os elementos Arsênio (As), Telúrio (Te), Chumbo (Pd) e Ouro (Au) e também da mesma forma, Caracelli (Caracelli, Zukerman-Schpector e Tiekink, 2012) identificou o Selênio (Se).

Responsáveis pela estrutura secundária das proteínas e também pela sua interação com ligantes e fármacos, estando atuando também na estabilização do DNA, o fenômeno que dá origem à essas interações baseia-se na nuvem eletrônica do anel aromático (Figura 1.6) atuando como doador de elétrons e assim, interagindo com a carga positiva do cátion (Ma e Dougherty, 1997; Gallivan e Dougherty, 1999; Wintjens et al., 2000; Michael Gromiha et al., 2004; Dougherty, 2007; Du, Meng, et al., 2012; Pinheiro et al., 2017).



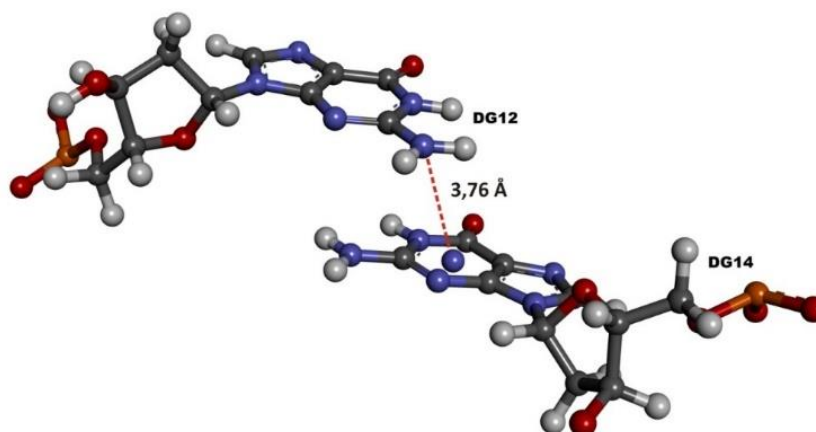
**Figura 1.6** Nuvem eletrônica de um anel aromático que atua como doadora de elétrons para uma interação cátion $\cdots\pi$ , sendo a região eletronegativa de coloração vermelha. Extraído de Dougherty (2007).

A representação esquemática de uma interação cátion $\cdots\pi$  é mostrada na Figura 1.7.

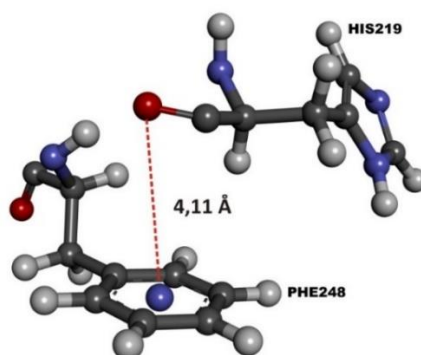


**Figura 1.7** Esquema de uma interação cátion $\cdots\pi$ , representando-se um cátion genérico posicionado sobre a face de um anel de benzeno (Pinheiro et al., 2017).

Segundo Zhang (Zhang et al., 2011) e colaboradores as interações N $\cdots\pi$  (Figura 1.8) e O $\cdots\pi$  (Figura 1.9) também são consideradas como interações Cátion $\cdots\pi$ , pois nesses casos, o nitrogênio e oxigênio são elétrons deficientes.



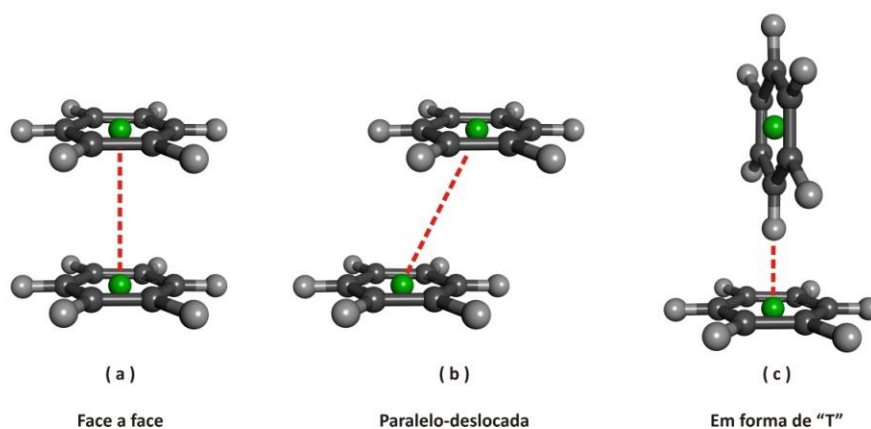
**Figura 1.8** Representação gráfica de uma interação N $\cdots\pi$  entre as bases G 12 e G 14 de uma molécula de DNA, imagem gerada a partir de dados cristalográficos (Drew et al., 1981) e com a utilização do software Discovery Studio 3.1<sup>TM</sup>, extraído de Sacco (2013).



**Figura 1.9** Representação gráfica de uma interação  $O \cdots \pi$  entre a HIS219 e a PHE248 de glutathione reductase, imagem gerada a partir de dados cristalográficos (Karplus e Schulz, 1989) e visualizada com a utilização do software Discovery Studio 3.1™, extraído de Sacco (2013).

### 1.1.5 INTERAÇÕES $\pi \cdots \pi$

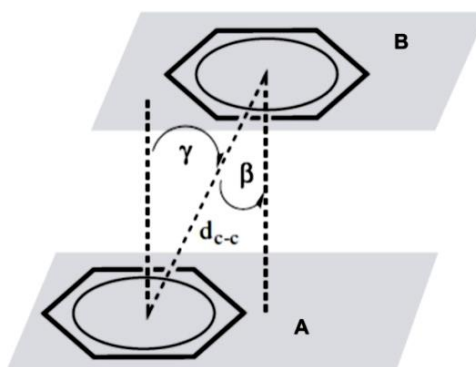
Embora a literatura (Meyer et al., 2003) refira-se a três tipos de interações  $\pi$ , sendo estas as interações  $\pi$  face a face, interações  $\pi$  paralelas-deslocadas e as interações  $\pi$  em forma de “T”, conforme ilustrado na Figura 1.10, podendo-se considerar que as interações  $\pi$  em forma de “T” são na realidade uma interação  $C-H \cdots \pi$  (Levitt e Perutz, 1988; Cunha et al., 2006; Birchall et al., 2011), sendo que no caso desta última, o carbono ligado ao hidrogênio não pertence à um anel, Meyer et al. (2003).



**Figura 1.10** Representações gráficas dos tipos de interação  $\pi \cdots \pi$ : **(a)** face a face, **(b)** paralela-deslocada, **(c)** forma de “T”. Adaptado de Singh (Singh et al., 2009) e Sherrill (Smith et al., 2008; Sherrill et al., 2009).

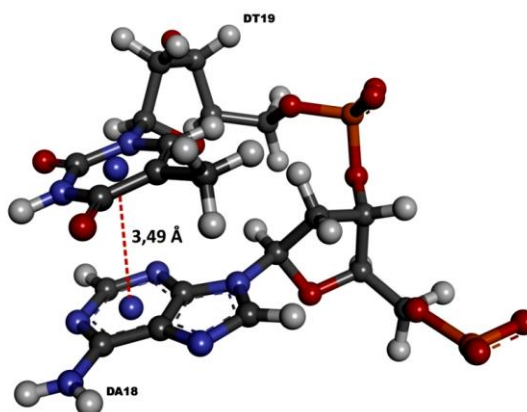
Conhecendo essas informações, devem ser consideradas que para o estudo das interações  $\pi$  face a face e face-deslocada, são utilizados os parâmetros geométricos esquematizados na Figura 1.11 (Janiak, 2000; Meyer et al., 2003; Tomić et al., 2006).

- $d_{c-c}$  : Distância existente entre os centroides dos anéis aromáticos **A** e **B**;
- $\gamma$  : ângulo existente entre o vetor normal ao plano do anel **A** e  $d_{c-c}$ ;
- $\beta$  : ângulo existente entre o vetor normal ao plano do anel **B** e  $d_{c-c}$ .



**Figura 1.11** Parâmetros geométricos para as interações  $\pi\cdots\pi$ . Adaptado de Tomić et al. (2006).

A Figura 1.12 mostra um exemplo de uma interação  $\pi\cdots\pi$  em meio biológico, segundo pesquisa efetuada pelo autor para a sua dissertação de mestrado (Sacco, 2013).



**Figura 1.12** Representação gráfica de uma interação  $\pi\cdots\pi$  entre as bases DA 18 e DT 19 de uma molécula com a imagem gerada a partir de dados cristalográficos (Drew et al., 1981) e com a utilização do software Discovery Studio 3.1™, extraído de Sacco (2013)

### 1.1.6 PARÂMETROS GEOMÉTRICOS DAS INTERAÇÕES $\pi$

Como resultado da sua pesquisa, Sacco (2013) obteve em trabalhos de diversos pesquisadores os parâmetros geométricos teóricos para as principais interações  $\pi$  abordadas nos itens 1.1.2, 1.1.3, 1.1.4 e 1.1.5, os quais são agrupados na Tabela 1.2.

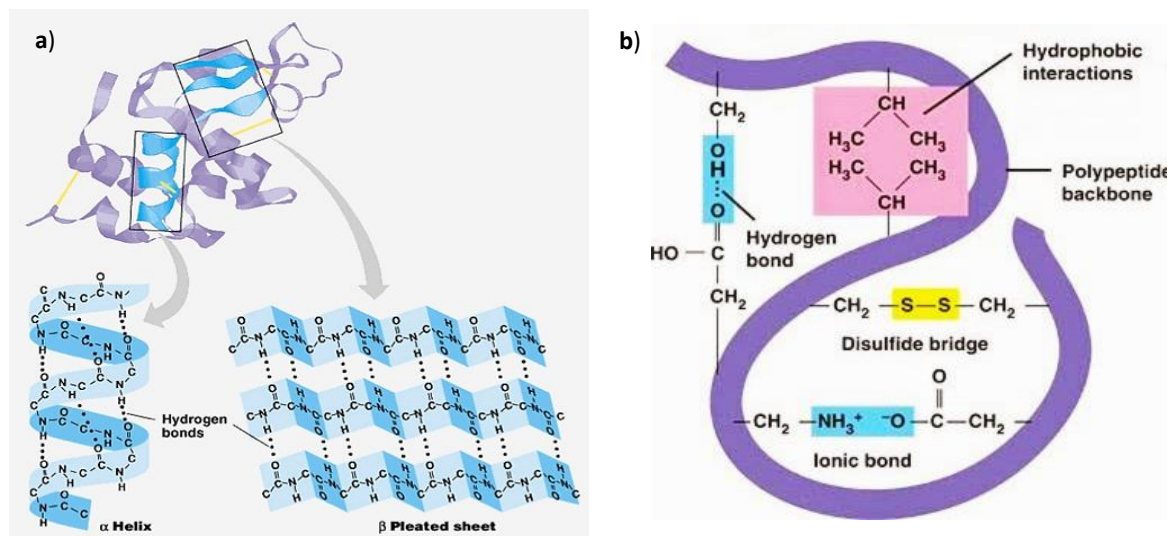
**Tabela 1.2** Parâmetros geométricos teóricos para as interações  $\pi$ , por tipo, segundo Sacco (2013)

Tipo de Interação	Módulo do vetor distância (Å)	Ângulo entre o vetor normal e o vetor distância (°)
$\pi \cdots \pi$	$3,0 \leq d \leq 4,5$	$\alpha$ e $\beta \leq 30$
C-H $\cdots\pi$	$2,5 \leq d \leq 3,8$	$\alpha \leq 30$
N-H $\cdots\pi$	$2,5 \leq d \leq 3,8$	$\alpha \leq 30$
O-H $\cdots\pi$	$2,5 \leq d \leq 3,8$	$\alpha \leq 30$
N $\cdots\pi$	$3,2 \leq d \leq 4,5$	$\alpha \leq 30$
O $\cdots\pi$	$3,2 \leq d \leq 4,5$	$\alpha \leq 30$

Dessa forma, para que uma interação  $\pi$  possa ocorrer, a localização dos átomos ou centroides com relação ao plano do anel de onde tem origem essa interação não covalente, deve necessariamente atender aos parâmetros estipulados na tabela acima.

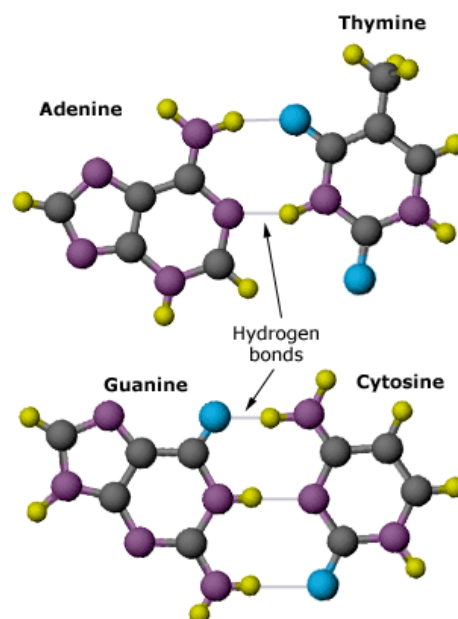
## 1.2. LIGAÇÕES DE HIDROGÊNIO

A ligação de hidrogênio é uma interação fundamental para a conformação da estrutura do DNA e as estruturas secundárias e terciárias de proteínas (Figura 1.13), contribuindo para que estas sejam estabilizadas (Pauling et al., 1951; Kollman e Allen, 1972; Steiner e Koellner, 2001).



**Figura 1.13** Representação esquemática das ligações de hidrogênio nas estruturas secundária (a) e terciária (b) das proteínas, extraído de BIT 603 (Caracelli e Zukerman-Schpector, 2010).

É importante ressaltar que estas interações não covalentes tornaram-se notórias após a publicação do trabalho *Molecular structure of nucleic acids* (Watson e Crick, 1953) onde pela primeira vez foi descrito o seu papel funcional nas estruturas de DNA e devido a isso, conforme denominado por Kool (Kool, 2001), são conhecidas como *Watson-Crick hydrogen bonding* e representadas na Figura 1.14.



**Figura 1.14** Representação esquemática das *Watson-Crick hydrogen bonding* nas ligações entre fitas do DNA efetuada pelos pares de bases G-C e A-T, (Lindberg et al., 2017).

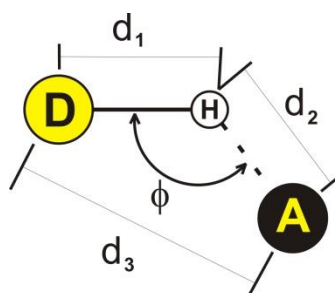
Segundo Ermondi e Caron (2006), estas interações ocorrem quando um átomo (acceptor) interage com um átomo de hidrogênio ligado covalentemente à um segundo átomo eletronegativo (doador), sendo essa interação potencializada em função das dimensões do átomo de hidrogênio que permite a proximidade da interação entre dipolos e cargas. Segundo Steiner (Steiner, 2002), considerando a configuração  $X-H\cdots A$ , o par X-H é o acceptor de elétron e A o doador, a Tabela 1.3 mostra ambos os grupos.

**Tabela 1.3** Acceptores e doadores de elétrons em ligações de hidrogênio, (Aakeroy e Seddon, 1993)

Acceptor de elétron X-H	Doador de elétron A
C-H	C = C, C $\equiv$ C, arenes
N-H	N
P-H	P
O-H	O
S-H	F
F-H	S
Cl-H	Cl
Br-H	Br
I-H	I

### 1.2.1 PARÂMETROS GEOMÉTRICOS DAS LIGAÇÕES DE HIDROGÊNIO

Tal qual as interações  $\pi$ , essas interações não covalentes são dependentes de parâmetros geométricos específicos (Figura 1.15).



**Figura 1.15** Representação esquemática de uma ligação de hidrogênio, considerando como doador o átomo ligado ao hidrogênio e acceptor um segundo átomo ou centroide. Extraído de Ermondi e Caron (2006).

**Tabela 1.4** Parâmetros geométricos das ligações de hidrogênio representados na Figura 1.13, adaptado de Ermondi e Caron (2006)

Parâmetro	Valor
$d_3$	$2,5\text{Å} < d_3 < 3,2\text{Å}$
$\phi$	$130^\circ < \phi < 180^\circ$

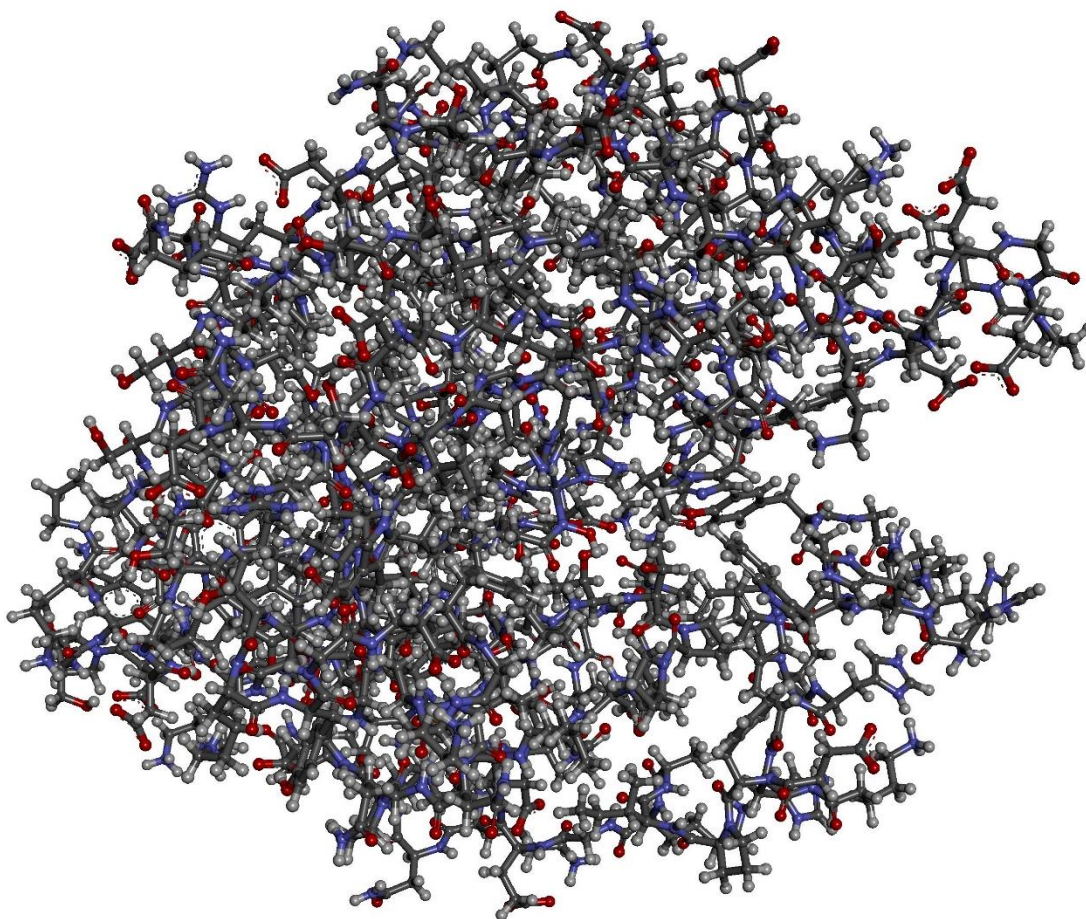
Os valores dos demais parâmetros, são funções de valores relacionados ao comprimento da ligação covalente ( $d_1$ ) e ( $d_2$ ) de distância próximo à soma dos raios de van der Waals, (Kollman e Allen, 1972). O entendimento destes parâmetros geométricos será útil para a compreensão dos conceitos envolvidos no Mapeamento Vetorial.

### 1.3 O MAPEAMENTO VETORIAL

O conceito do Mapeamento Vetorial é decorrente de uma demanda identificada na resolução de um exercício da disciplina BIT 603 (Caracelli e Zukerman-Schpector, 2009), onde, utilizando o visualizador gráfico Discovery Studio 3.1<sup>TM</sup> ( Accelrys, 2012), deveriam ser identificadas em tela gráfica as interações  $\pi$  existentes em uma estrutura cristalográfica e obter para estas, as informações dos parâmetros identificados anteriormente nas Figuras 1.1 e 1.11. Embora a biomolécula selecionada possuísse poucos átomos, a sua visualização na tela de um monitor exibiu uma imagem saturada (Figura 1.16), fazendo com que o processo de busca fosse trabalhoso e sujeito a erro humano. A necessidade de melhoria desse processo mostrou ser um fato instigante e algumas semanas depois, a perspectiva da utilização de conceitos de Álgebra Linear, especificamente operações com vetores, passaram a ser consideradas promissoras. Com o conhecimento de que as estruturas cristalográficas depositadas em bancos de dados (Embl-Ebi, 1995; Berman *et al.*, 2000; Berman *et al.*, 2001; Berman *et al.*, 2007; Jersey, 2012) possuem a característica de apresentar a posição de seus átomos por meio de coordenadas espaciais onde a posição



de um átomo é  $P(x_p; y_p; z_p)$ , facilitou a utilização do cálculo vetorial para o processamento dos parâmetros geométricos dessas interações não covalentes.



**Figura 1.16** Representação gráfica uma enzima, cuja imagem foi gerada a partir de dados cristalográficos (Scolnick et al., 1997), utilizando o software Discovery Studio 3.1™.

Um exemplo de um arquivo disponível no PDB, porem com o formato do seu arquivo alterado para mol2<sup>1</sup>, onde são disponibilizadas as coordenadas  $x$ ,  $y$  e  $z$  de cada

---

<sup>1</sup> Os arquivos disponíveis no PDB possuem a extensão “.pdb”, em função da conveniência do usuário e com a utilização de visualizadores gráficos, podem ser alterados para outros formatos (mol2, dsv, tif, jpg, dentre outros).

átomo pode ser observada na Tabela 1.5, esse tipo de padrão possibilitou a evolução do conceito do Mapeamento Vetorial para a determinação de interações não covalentes e dependentes de parâmetros geométricos (Sacco, 2013).

**Tabela 1.5** Coordenadas espaciais da Tyr 23 existente na estrutura cristalográfica de uma glutiona redutase (Karplus e Schulz, 1989), adaptado de Sacco (2013)

Posição do átomo na estrutura	Átomo	Coordenada X	Coordenada Y	Coordenada Z	Tipologia	Posição do aminoácido na estrutura	Aminoácido
74	N	46,0470	44,7800	33,5440	N.am	6	TYR23
75	CA	47,2170	45,4710	33,0010	C.3	6	TYR23
76	C	46,8840	45,8860	31,6060	C.2	6	TYR23
77	O	46,4910	45,0560	30,8030	O.2	6	TYR23
78	CB	48,3240	44,4240	32,9080	C.3	6	TYR23
79	CG	49,7220	44,9920	32,8250	C.ar	6	TYR23
80	CD1	50,0640	46,1810	33,4690	C.ar	6	TYR23
81	CD2	50,7180	44,2830	32,1690	C.ar	6	TYR23
82	CE1	51,3700	46,6650	33,4530	C.ar	6	TYR23
83	CE2	52,0300	44,7410	32,1490	C.ar	6	TYR23
84	CZ	52,3390	45,9480	32,7660	C.ar	6	TYR23
85	OH	53,6390	46,4000	32,7500	O.3	6	TYR23
86	HD2	50,4690	43,3577	31,6644	H	6	TYR23
87	HE1	51,6243	47,5837	33,9671	H	6	TYR23
88	HE2	52,8036	44,1637	31,6579	H	6	TYR23
89	HH	53,9814	46,4455	33,6703	H	6	TYR23
90	HA	47,5048	46,3226	33,6050	H	6	TYR23
91	HB1	48,1494	43,8256	32,0224	H	6	TYR23
92	HB2	48,2672	43,7935	33,7867	H	6	TYR23
93	HD1	49,2980	46,7396	33,9925	H	6	TYR23
94	H	45,4369	44,2711	32,9220	H	6	TYR23

### 1.3.1 OPERAÇÕES COM VETORES

Os vetores são muito utilizados na resolução de cálculos nas Ciências Exatas, pois estão associados a grandezas que necessitam de módulo, direção e sentido para serem completamente especificadas (Menon, 2009). Dessa forma, as operações com vetores são utilizadas pela Álgebra Linear para equacionar problemas trigonométricos associados a essas grandezas. No caso do estudo das biomoléculas pode ser utilizado como base de um algoritmo para localizar e identificar tanto as interações  $\pi$  ou as ligações de hidrogênio em função de variáveis de contorno ou características espaciais, conforme exposto nas Tabelas 1.2 e 1.4.

### 1.3.2 O CÁLCULO VETORIAL APLICADO ÀS BIOMOLÉCULAS

Nos próximos dois itens, serão mostrados de forma abreviada como utilizar o cálculo vetorial para localizar e identificar as interações não covalentes abordadas nessa pesquisa.

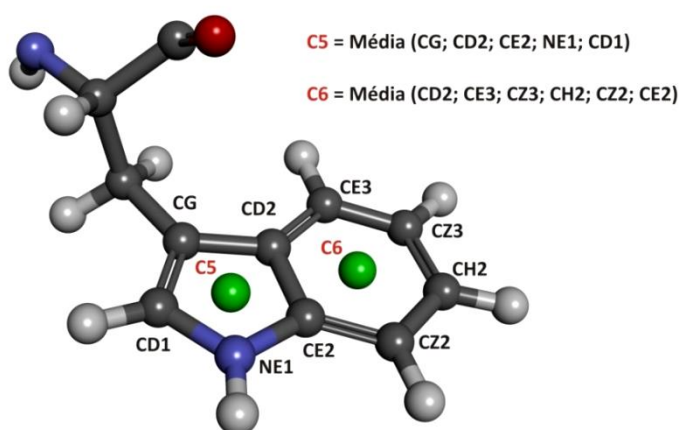
#### 1.3.2.1 INTERAÇÕES $\pi$

Considerando que uma interação  $\pi$  tem origem no centroide de anéis aromáticos, conforme abordado nos itens 1.1.1 a 1.1.6, este centroide deve ser conhecido e a sua coordenada  $C(x_c; y_c; z_c)$ , calculada. Como já é de conhecimento que os átomos de uma estrutura cristalográfica são representados por coordenadas, essa tarefa pode ser executada da forma a seguir:

$$C(x_c; y_c; z_c) = [P_1(x_1; y_1; z_1) + P_2(x_2; y_2; z_2) + P_3(x_3; y_3; z_3) + \dots + P_n(x_n; y_n; z_n)] / n \quad (1.1)$$

$$\text{Onde: } x_c = (x_1 + x_2 + x_3 + \dots + x_n) / n; \quad y_c = (y_1 + y_2 + y_3 + \dots + y_n) / n \quad \text{e} \quad z_c = (z_1 + z_2 + z_3 + \dots + z_n) / n$$

Conforme a Equação 1.1,  $C$  é a coordenada do centroide resultante da média aritmética das coordenadas  $x_n$ ,  $y_n$  e  $z_n$ , podendo os valores de  $n$  variarem de 5 ou mais átomos, conforme exemplificado na Figura 1.15.



**Figura 1.17** Localização dos átomos envolvidos no cálculo dos centros de um triptofano. Imagem obtida a partir de estrutura cristalográfica (Karplus e Schulz, 1989) visualizada utilizando o Discovery Studio 3.1<sup>TM</sup>. Extraído de Sacco (2013).

Uma vez conhecida a coordenada do centroide, utiliza-se esta coordenada como sendo um ponto fixo e utiliza-se a Equação 1.2 para calcular as distâncias do centroide a todos os demais átomos e centros de biomolécula.

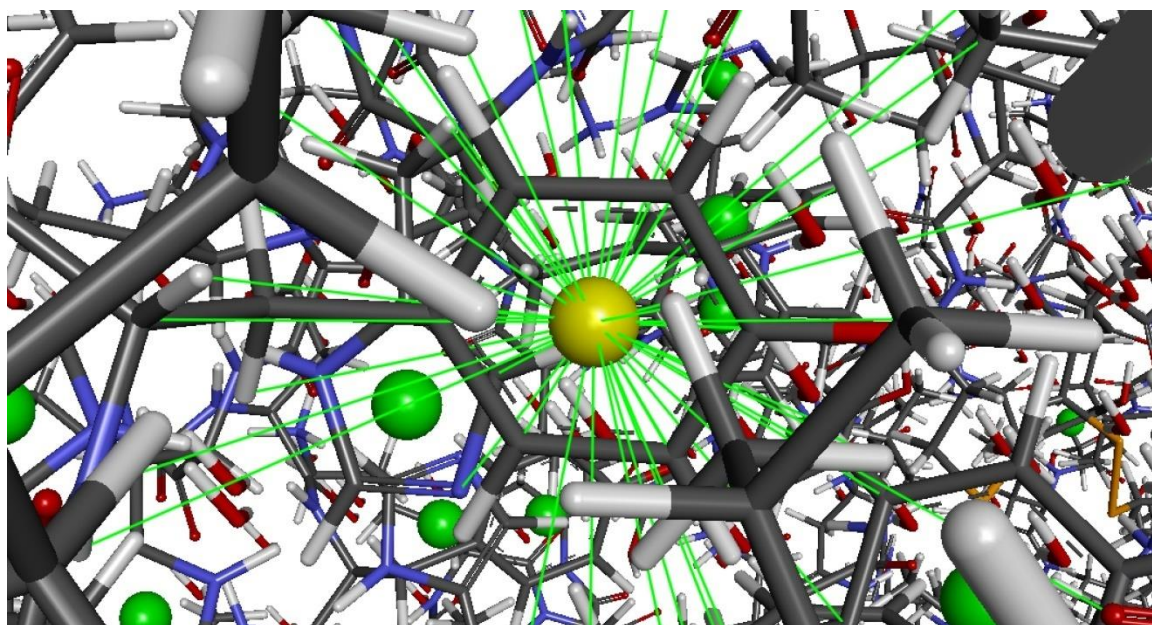
$$\vec{v}(x; y; z) = \overrightarrow{AB} = (x_2; y_2; z_2) - (x_1; y_1; z_1)$$

$$\therefore \vec{v} = \overrightarrow{AB} = (x_2 - x_1; y_2 - y_1; z_2 - z_1) \quad (2.2)$$

Assim sendo, um vetor<sup>2</sup>  $\vec{v} = \overrightarrow{AB}$ , possui a sua origem no centroide e extremidade em um átomo ou centroide da estrutura

$$\begin{aligned}\vec{v}(x; y; z) &= \overrightarrow{AB} = (x_2; y_2; z_2) - (x_1; y_1; z_1) \\ \therefore \vec{v} &= \overrightarrow{AB} = (x_2 - x_1; y_2 - y_1; z_2 - z_1)\end{aligned}\tag{1.2}$$

Caso fosse possível visualizar esses vetores com origem no centroide, a imagem seria semelhante à Figura 1.18.



**Figura 1.18** Representação gráfica dos vetores com origem no centroide de um anel com destino em átomos localizados à uma distância máxima de 4,5 Å. Visão parcial da estrutura de uma glutatona redutase gerada a partir de dados cristalográficos (Karplus e Schulz 1989), e visualizada com a utilização do software Discovery Studio 3.1<sup>TM</sup>.

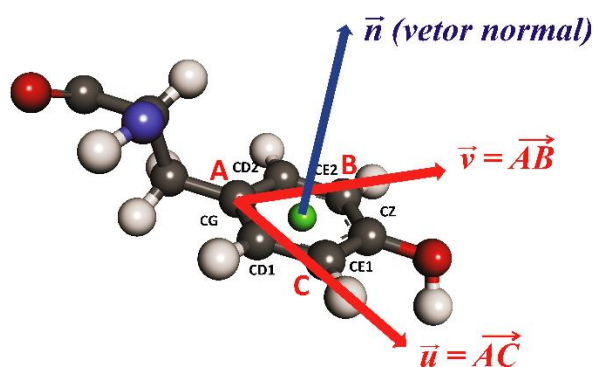
---

<sup>2</sup> Definições detalhadas sobre vetores encontradas em livros de álgebra linear. Por exemplo, veja Steinbruch (1987) e Leithold (1994).

Todos os vetores obtidos para devem ser armazenados para que possam ser calculadas o comprimento das interações  $\pi$ . Conforme a Tabela 1.2, em atenção aos parâmetros geométricos para as interações  $\pi$ , interessam-nos distâncias entre o centroide de origem e demais coordenadas que atendam a  $2.5\text{\AA} \leq d \leq 4,5\text{\AA}$ . Para que essa distância seja obtida, utiliza-se a equação 1.3 a qual determina o módulo do vetor  $\vec{v} = \overrightarrow{AB}$ .

$$\|\vec{v}\| = \|\overrightarrow{AB}\| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \quad (1.3)$$

Os módulos dos vetores que atendem a  $2.5\text{\AA} \leq d \leq 4,5\text{\AA}$ , são preservados para cálculos posteriores e os demais descartados. Em continuidade, sabendo-se que para que ocorra uma interação entre um centroide e um átomo ou outro centroide, além destes últimos estarem à uma distância  $2.5\text{\AA} \leq d \leq 4,5\text{\AA}$ , o vetor  $\vec{v} = \overrightarrow{AB}$  deve formar com o vetor normal ao plano do anel um ângulo  $\alpha \leq 30^\circ$ . Considerando-se que três pontos quaisquer formam um plano, ao serem selecionadas as coordenadas de três átomos pertencentes ao anel aromático, é possível representar vetorialmente o plano desse anel (Figura 1.19).



**Figura 1.19** Determinação do vetor normal ao plano do anel aromático de resíduo de aminoácido tirosina. Imagem obtida a partir da estrutura cristalográfica (Karplus e Schulz, 1989) e visualizada utilizando o Discovery Studio 3.1™, extraído de Sacco (2013).

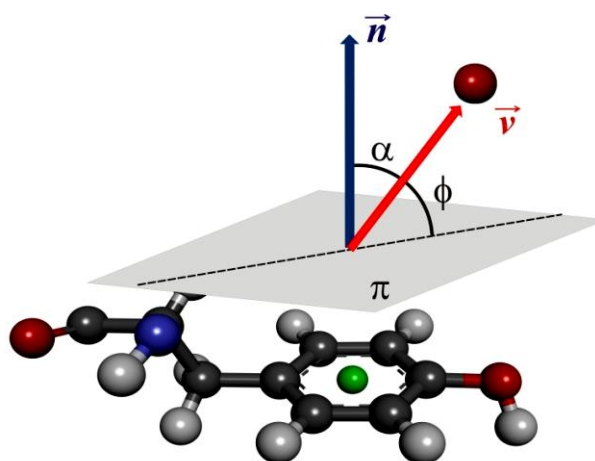
Conhecidas as coordenadas referentes aos pontos **A**, **B** e **C**, são obtidos os vetores  $\vec{v} = \overrightarrow{AB} = (x_1; y_1; z_1)$  e  $\vec{u} = \overrightarrow{AC} = (x_2; y_2; z_2)$  pertencentes ao plano do anel

aromático, o que permite o cálculo do vetor normal ao plano ( $\vec{n}$ ), o qual é o determinante de uma matriz quadrada de 3 colunas por 3 linhas, calculado pela Equação 1.4.

$$\vec{n} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{vmatrix}$$

$$\vec{n} = [(y_1 z_2) - (y_2 z_1)]\vec{i} + [(z_1 x_2) - (z_2 x_1)]\vec{j} + [(x_1 y_2) - (x_2 y_1)]\vec{k} \quad (1.4)$$

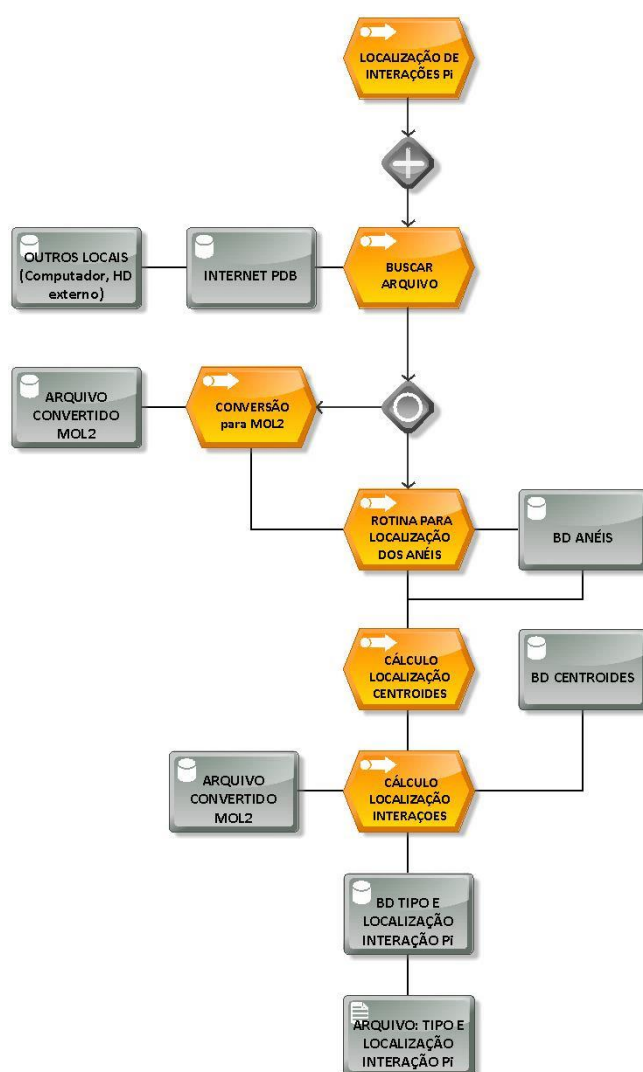
Considerando que cada centroide está situado no plano do anel aromático e perpendicular a este plano está o vetor  $\vec{n}$ , calculado pela Equação 1.4, como consequência, é obtido o ângulo  $\alpha$  (Figura 1.20) existente entre os vetores  $\vec{n}$  e  $\vec{v}$ , o qual é fornecido pela Equação 1.5 (Steinbruch, 1987).



**Figura 1.20** Determinação do ângulo  $\alpha$  entre o vetor normal  $\vec{n}$  ao plano  $\pi$  formado pelo anel aromático e o vetor distância  $\vec{v}$  com base na estrutura de uma tirosina. Imagem adaptada de Steinbruch (1987) sobre estrutura cristalográfica (Karplus e Schulz, 1989) e visualizada utilizando o Discovery Studio 3.1<sup>TM</sup>, extraído de Sacco (2013).

$$\sin \alpha = \frac{|\vec{v} \cdot \vec{n}|}{|\vec{v}| \cdot |\vec{n}|}, \text{ assim sendo, } \alpha = \sin^{-1} \frac{|\vec{v} \cdot \vec{n}|}{|\vec{v}| \cdot |\vec{n}|} \quad (1.5)$$

Uma vez reconhecida a importância dos cálculos referentes à determinação das coordenadas dos centroides, do módulo do vetor distância  $\vec{v}$ , do vetor normal ao plano  $\vec{n}$  e do ângulo  $\alpha$  entre estes vetores, fica evidente a necessidade de utilização de recursos computacionais para efetuar as todas as operações matemáticas necessárias, uma vez que a quantidade de combinações centroide-átomo e centroide-centroide afetam geometricamente o volume dos cálculos envolvidos, os quais em proteínas podem ultrapassar 14 mil átomos. Em sua pesquisa, Sacco (2013) utilizou com sucesso a planilha eletrônica Microsoft Excel (Microsoft.Com, 2012) para realizar de forma semiautomática os cálculos acima descritos, conforme o fluxograma da Figura 1.21.

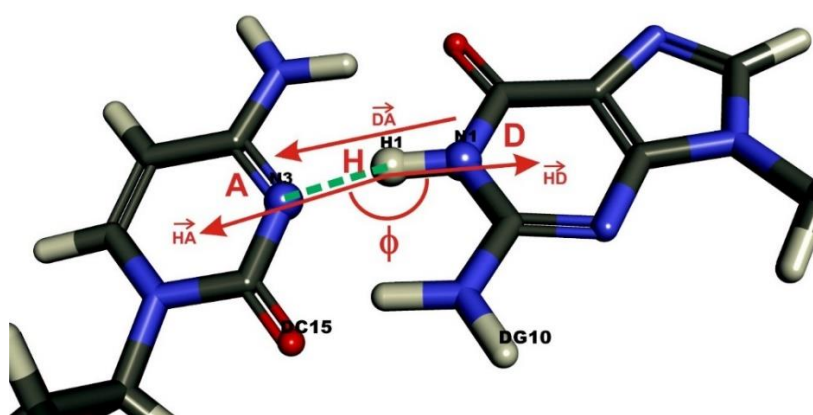


**Figura 1.21** Fluxograma das operações envolvidas no mapeamento vetorial para as interações  $\pi$ , processo modelado com a utilização do software ARIS (ARIS Express, 2009), extraído de Sacco (2013).



### 1.3.2.2 LIGAÇÕES DE HIDROGÊNIO

Considerando o pressuposto dos itens 1.2.1 e 1.2.2, também pode-se utilizar o mapeamento vetorial para localizar as ligações de hidrogênio, sendo este procedimento mais simples do que os cálculos envolvidos nas interações  $\pi$ . Para melhor compreensão, foi criada a Figura 1.22, onde é possível visualizar os vetores envolvidos nessa ligação não covalente.



**Figura 1.22** Representação dos vetores e o ângulo  $\phi$  entre estes, utilizados para a localização de ligações de hidrogênio utilizando o Mapeamento Vetorial. Imagem gerada a partir de estrutura cristalográfica (Teng et al., 1988) e visualizada utilizando o Discovery Studio 3.1<sup>TM</sup>, autoria própria.

Consonante à representação da Figura 1.15, a ilustração acima mostra os vetores  $\overrightarrow{HD}$  que representa a ligação covalente existente entre o hidrogênio e átomo doador, o vetor  $\overrightarrow{HA}$  a ligação de hidrogênio existente entre o aceitador e o átomo de hidrogênio, e o vetor  $\overrightarrow{DA}$  como sendo o parâmetro  $d_3$  da Tabela 1.4. Por analogia, sabe-se que  $\overrightarrow{HD}$  e  $\overrightarrow{HA}$  são os vetores que determinam o ângulo  $\phi$ , que é o segundo parâmetro da referida tabela. Assim, com base nas informações acima e utilizando a Equação 1.2, o resultado obtido é:

$$\overrightarrow{HD} = (x_D - x_H; y_D - y_H; z_D - z_H)$$

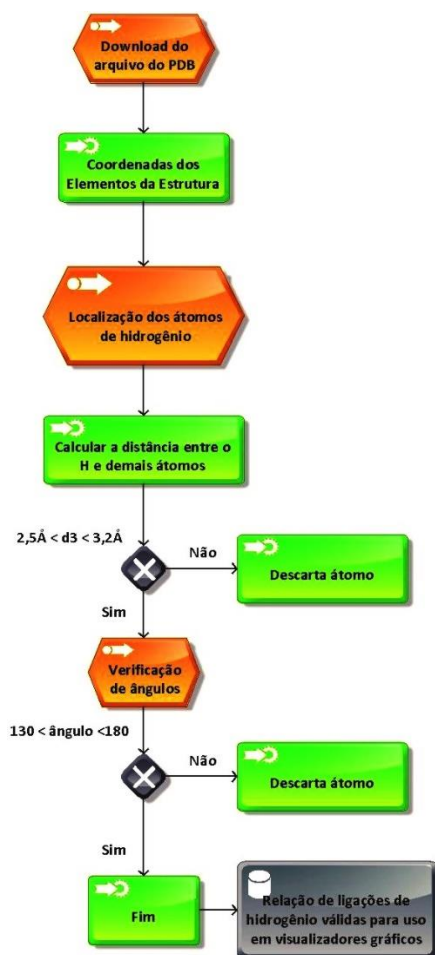
$$\overrightarrow{HA} = (x_A - x_H; y_A - y_H; z_A - z_H)$$

$$d_3 = \overrightarrow{DA} = (x_A - x_D; y_A - y_D; z_A - z_D)$$

Possuindo as informações acima, o ângulo  $\phi$  é calculado utilizando-se a Equação 1.5:

$$\sin \phi = \frac{|\overline{DA} \cdot \overline{HA}|}{|\overline{DA}| \cdot |\overline{HA}|}, \text{ assim sendo, } \phi = \sin^{-1} \frac{|\overline{DA} \cdot \overline{HA}|}{|\overline{DA}| \cdot |\overline{HA}|}$$

Dessa forma, a partir da localização dos átomos de hidrogênio e com o auxílio de ferramentas computacionais, é possível construir de forma análoga às interações  $\pi$ , um banco de dados com os valores  $d_3$  e  $\phi$  válidos para as ligações de hidrogênio. A Figura 1.21 mostra, de forma resumida, o macroprocesso idealizado para o Mapeamento Vetorial das ligações de hidrogênio.



**Figura 1.23** Fluxograma resumido das operações envolvidas no mapeamento vetorial para a localização das ligações de hidrogênio, macroprocesso modelado com a utilização do software ARIS (ARIS Express, 2009), elaboração própria.

# Capítulo 2

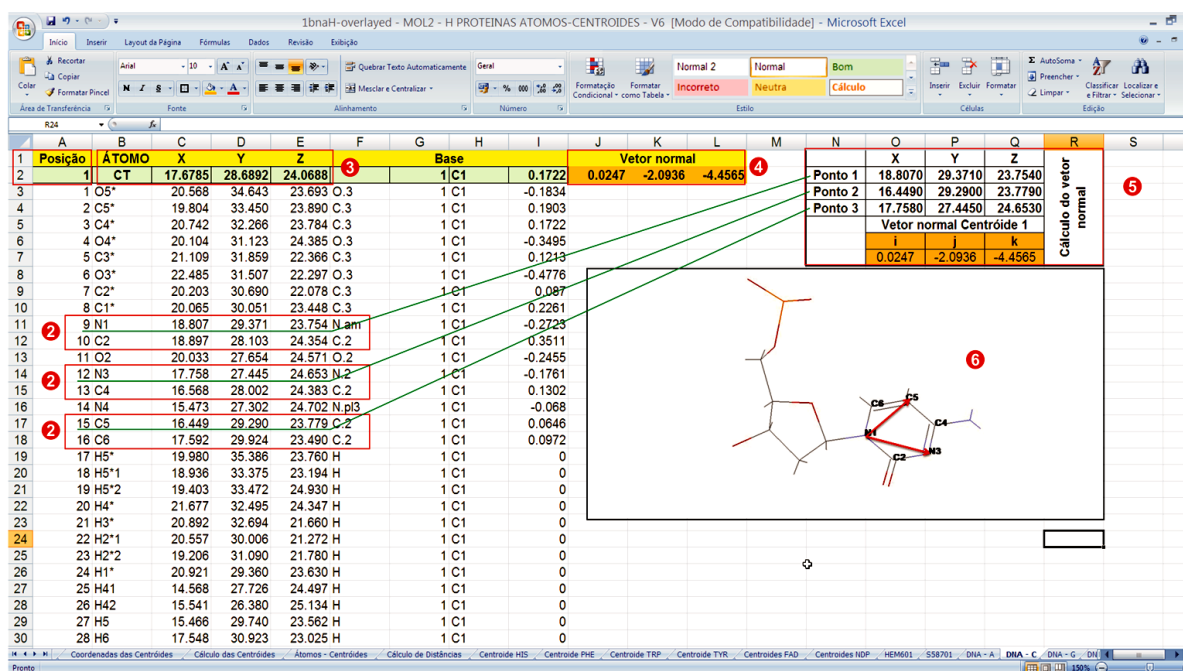
---

MATERIAIS E MÉTODOS

O objetivo deste capítulo é apresentar os métodos utilizados em que os conceitos matemáticos da álgebra vetorial são associados às coordenadas das estruturas moleculares obtidas por experimentos cristalográficos depositadas no PDB. Serão abordados os primeiros experimentos efetuados pelo autor utilizando planilha eletrônica e os seus resultados. Em função desses resultados obtidos, desenvolveu-se um software específico, automatizando o processamento do mapeamento vetorial. Os aspectos do desenvolvimento do software serão abordados nesse capítulo e também novos experimentos efetuados com relação às características das interações  $\pi$  e ligações de hidrogênio, obtendo uma nova visão sobre a conformação em um grupo específico de estruturas de DNA.

## **2 A APLICAÇÃO DO MAPEAMENTO VETORIAL – O INÍCIO**

A oportunidade de inovação identificada por meio do exercício da disciplina BIT 603 (Caracelli e Zukerman-Schpector, 2009), evoluiu como pesquisa de uma dissertação de mestrado intitulada “Mapeamento vetorial de interações eletrostáticas para biomoléculas” (Sacco, 2013), cujo propósito é a localização e identificação de interações  $\pi$ , aplicando os conceitos expostos nos itens 1.3.1 e 1.3.2 do Capítulo 1. Dessa forma, criou-se em um planilha eletrônica do Excel (Microsoft.Com, 2012a; b) uma rotina utilizando programação com macro comandos para automatizar algumas rotinas, tais como o cálculo para a determinação simultânea das coordenadas dos centroides e vetor normal para os anéis aromáticos selecionados (Figura 2.1), essa primeira etapa do mapeamento, servia como base para o cálculo das interações  $\pi$  válidas. Importante salientar que a planilha eletrônica tem como principal função a execução dos cálculos vetoriais, sendo que a seleção dos anéis aromáticos é feita manualmente. Todavia, embora parcialmente automatizado, esse procedimento mostrou ser mais rápido e confiável do que os procedimentos anteriormente adotados.



**Figura 2.1** Imagem da tela da planilha eletrônica construída para o cálculo da coordenada do centroide e determinação do vetor normal ao plano do anel de uma base Citosina, adaptado de Sacco (2013), onde:

- 1- Posição da base na estrutura do DNA;
- 2- Os átomos cujas coordenadas estão envolvidas no cálculo do centroide;
- 3- As células "C2", "D2" e "E2" da planilha contém a Equação 1.1 e expressam o resultado do cálculo das coordenadas  $x$ ,  $y$  e  $z$  do centroide;
- 4- Versores  $i$ ,  $j$  e  $k$  do vetor normal ao plano do anel;
- 5- Células "O7", "P7" e "Q7" da planilha são utilizadas para o cálculo da Equação 1.4 processando os valores das coordenadas  $x$ ,  $y$  e  $z$  dos três pontos selecionados no anel aromático e expressam os valores dos versores  $i$ ,  $j$  e  $k$ ;
- 6- Esquema da base identificando os pontos cujas coordenadas são utilizadas nos cálculos. Planilha eletrônica do Microsoft Excel.

A execução da rotina acima exposta para todos os anéis aromáticos da biomolécula gera uma segunda planilha que inicialmente (Figura 2.2) possuía o propósito de ser a origem de uma possível interação  $\pi$ , sendo utilizadas para o cálculo simultâneo das distâncias deste centroide de origem até os demais átomos e outros centroides (cálculo do módulo do vetor distância) e cálculo do ângulo  $\alpha$  (Figura 2.3). Os valores de  $d$  e  $\alpha$  que estiverem dentro dos parâmetros previstos na Tabela 1.2 caracterizam interações  $\pi$  válidas, porém, essas interações consideradas válidas são selecionadas manualmente e

salvas como uma planilha Excel ou em formato mol2 para uso em visualizadores gráficos, como por exemplo o Discovery Studio 3.1™ (Accelrys, 2012).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1		Posição	ÁTOMO	X	Y	Z		Base			Vetor normal			
2	1	1	CT	16.9860	28.9525	24.7440	N,am	1	DC1	0.0000	0.0454	-2.2596	-4.2639	
3	2	2	CT	19.0686	26.6630	21.8522	N,am	2	DG2	0.0000	0.3194	1.0450	2.6454	
4	3	2	CT	18.2408	24.9478	22.6538	N,am	2	DG2	0.0000	0.4479	1.8701	4.4195	
5	4	3	CT	19.2200	23.7527	19.2010	N,am	3	DC3	0.0000	-0.4910	-2.1352	-4.2748	
6	5	4	CT	18.9680	20.3726	16.5868	N,am	4	DG4	0.0000	0.9462	0.9396	2.5163	
7	6	4	CT	17.0217	20.5253	17.2317	N,am	4	DG4	0.0000	1.4840	1.5724	4.2720	
8	7	5	CT	16.8670	17.8856	14.3842	N,am	5	DA5	0.0000	-0.8445	-0.9147	-2.6742	
9	8	5	CT	15.3767	19.3273	14.3915	N,am	5	DA5	0.0000	-1.4510	-1.4682	-4.5489	
10	9	6	CT	13.9228	16.9726	11.6410	N,am	6	DA6	0.0000	-0.8573	-0.7542	-2.7267	
11	10	6	CT	13.5612	18.9813	11.2567	N,am	6	DA6	0.0000	-1.3987	-1.0576	-4.6668	
12	11	7	CT	12.3068	17.6300	8.2173	N,am	7	DT7	0.0000	-1.3687	-0.5625	-4.6095	
13	12	8	CT	10.8322	18.9690	4.6475	N,am	8	DT8	0.0000	-0.7888	-0.2725	-4.7910	
14	13	9	CT	10.8145	20.7042	1.0197	N,am	9	DC9	0.0000	-0.5985	-0.1539	-4.7985	
15	14	10	CT	10.5042	21.6452	-2.9004	N,am	10	DG10	0.0000	0.4260	0.6080	2.7439	
16	15	10	CT	12.1765	20.4530	-2.8932	N,am	10	DG10	0.0000	0.7283	1.0072	4.5993	
17	16	11	CT	13.5883	22.1867	-6.7460	N,am	11	DC11	0.0000	0.7297	-0.3468	-4.7379	
18	17	12	CT	15.3144	21.2526	-10.2524	N,am	12	DG12	0.0000	0.0478	0.6095	2.7860	
19	18	12	CT	14.7003	19.3220	-9.8345	N,am	12	DG12	0.0000	0.1866	0.8898	4.7287	
20	19	13	CT	10.2607	15.7842	-9.1385	N,am	13	DC13	0.0000	-0.3178	0.6288	4.7673	
21	20	14	CT	13.1756	14.9480	-5.4686	N,am	14	DG14	0.0000	-0.5288	-0.3379	-2.7817	
22	21	14	CT	14.0548	16.7860	-5.8425	N,am	14	DG14	0.0000	-0.7753	-0.5364	-4.7525	
23	22	15	CT	14.8450	15.7928	-2.1177	N,am	15	DC15	0.0000	1.1639	1.1775	4.4991	
24	23	16	CT	17.5184	17.7462	0.8080	N,am	16	DG16	0.0000	-0.2423	-1.1447	-2.6165	
25	24	16	CT	15.9770	19.0583	0.4088	N,am	16	DG16	0.0000	-0.4447	-1.7324	-4.4493	
26	25	17	CT	17.7346	21.0458	3.1930	N,am	17	DA17	0.0000	0.1938	1.0724	2.7326	
27	26	17	CT	15.6648	21.1295	3.2950	N,am	17	DA17	0.0000	0.3088	1.7837	4.6814	
28	27	18	CT	16.5374	23.3292	6.1030	N,am	18	DA18	0.0000	0.1348	0.9852	2.7812	
29	28	18	CT	14.8987	22.1570	6.5993	N,am	18	DA18	0.0000	0.2316	1.6529	4.6910	
30	29	19	CT	14.7055	24.1038	9.4935	N,am	19	DT19	0.0000	-0.1742	1.5251	4.5938	
31	30	20	CT	13.0163	24.0010	13.2120	N,am	20	DT20	0.0000	0.0292	1.1504	4.6932	
32	31	21	CT	12.4320	23.2598	17.0342	N,am	21	DC21	0.0000	0.5583	1.0761	4.6607	
33	32	22	CT	12.1062	23.2692	20.9998	N,am	22	DG22	0.0000	-0.6646	-1.1290	-2.5361	
34	33	22	CT	14.0750	22.7160	20.7303	N,am	22	DG22	0.0000	-1.0942	-1.8432	-4.2971	
35	34	23	CT	14.0063	22.2200	24.7978	N,am	23	DC23	0.0000	1.3709	1.4190	4.4089	
36	35	24	CT	16.2776	22.4276	28.0540	N,am	24	DG24	0.0000	-0.7453	-1.1467	-2.4950	
37	36	24	CT	17.0698	24.0698	27.0993	N,am	24	DG24	0.0000	-1.0163	-1.8723	-4.2994	

**Figura 2.2** Planilha com a relação das coordenadas dos centroides e vetores normais ao plano de cada anel aromático calculados para o DNA – SL. Utilizada planilha eletrônica do Microsoft Excel.

A Figura 2.4 mostra uma interação válida, no caso uma interação  $N\cdots\pi$  obtida por meio de cálculo vetorial e usando uma planilha como gabarito, os dados espaciais obtidos são salvos, também manualmente, em um arquivo mol2.

Pos.	X	Y	Z	N.am	Base	DC	i	j	k
1 CT	16,9860	28,9525	24,7440	N.am	1 DC1	0,0000	0,04544	-2,25959	-4,26389
9 N1	18,0700	29,6610	24,3800	N.am	1 DC1	0,0000	#DIV/0!	90,0	90,0
13 C4	15,9170	28,2260	25,1200	C.2	1 DC1	0,0000	89,9	89,9	90,0
12 N3	17,1430	27,7610	25,3770	N.2	1 DC1	0,0000	89,8	89,8	90,0
10 C2	18,2240	28,4540	25,0150	C.2	1 DC1	0,0000	89,8	89,8	90,0
16 C6	16,8430	30,1710	24,1010	C.2	1 DC1	0,0000	2,49	-89,8	89,8
15 C5	15,7190	29,4420	24,4710	C.2	1 DC1	0,0000	2,49	90,0	90,0
28 H6	16,7403	31,1437	23,5912	H	1 DC1	0,0000	2,60	-89,8	89,8
27 H5	14,7030	29,8146	24,2575	H	1 DC1	0,0000	2,71	89,9	89,9
11 O2	19,3600	28,0140	25,2140	O.2	1 DC1	0,0000	2,83	89,0	89,0
14 N4	14,8280	27,4770	25,4440	N.pl3	1 DC1	0,0000	3,42	88,9	88,9
8 C1'	19,2310	30,4820	23,9440	C.3	1 DC1	0,0000		89,8	89,8

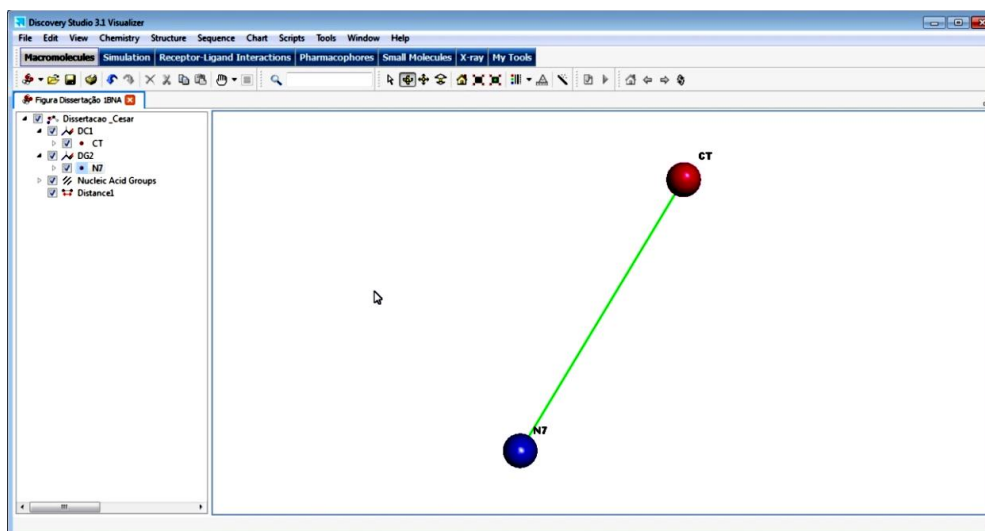
**Figura 2.3** Imagem da planilha eletrônica elaborada para o cálculo do módulo do vetor distância e os ângulos entre o vetor distância e o vetor normal para cada relação centroide-átomo e centroide-centroide da estrutura em estudo (Sacco, 2013):

- 1- Identificação do *array* do centroide no banco de dados de coordenadas e vetores normais;
- 2- Dados importados do banco de dados em função da identificação do *array*;
- 3- Estrutura determinada cristalograficamente com a identificação dos átomos e os centroides calculados já inseridos, com as coordenadas *x*, *y* e *z* de cada elemento;
- 4- Módulo do vetor distância relativo às coordenadas do centroide selecionado, obtido conforme a Equação 2.3, estando os resultados já ordenados do menor para o maior valor;
- 5- Ângulo entre o vetor normal ao plano e o vetor distância da coordenada do elemento relativa a coordenada do centroide, utilizada a Equação 2.5;
- 6- Módulo do ângulo entre o vetor normal ao plano e o vetor distância. Planilha eletrônica Microsoft Excel.

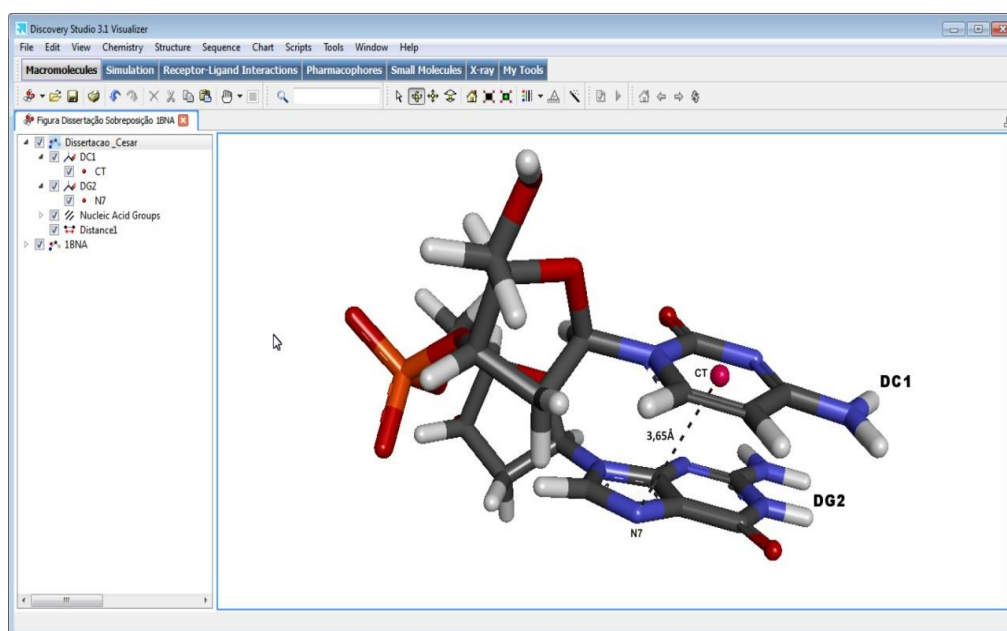
#	Created by:	Discovery	Studio						
#	Creation time:	Wed	October	15	16:36:20 E.	South	America	Standard	Time
@<TRIPOS>MOLECULE	Dissertacao Cesar	863	25	0	0				
@<TRIPOS>ATOM									
1 CT	16,9860	28,9525	24,7440	CT	1	DC1	0,0000		
42 N7	18,2070	27,4550	21,6380	N.ar	2	DG2	0,0000		

**Figura 2.4** Imagem da planilha eletrônica utilizada para exportar os resultados válidos uma interação  $\pi$ , na em arquivo compatível para visualização em software gráfico. Planilha eletrônica do Microsoft Excel (Sacco, 2013).

O arquivo mol2 da interação  $N\cdots\pi$  obtida por meio de cálculo vetorial pode ser visualizado em software gráfico (Figura 2.5) e utilizando o mesmo recurso computacional, Discovery Studio 3.1 (Inc., 2012), pode-se sobrepor a interação mapeada vetorialmente sobre o arquivo original, o qual foi utilizado para o mapeamento vetorial.



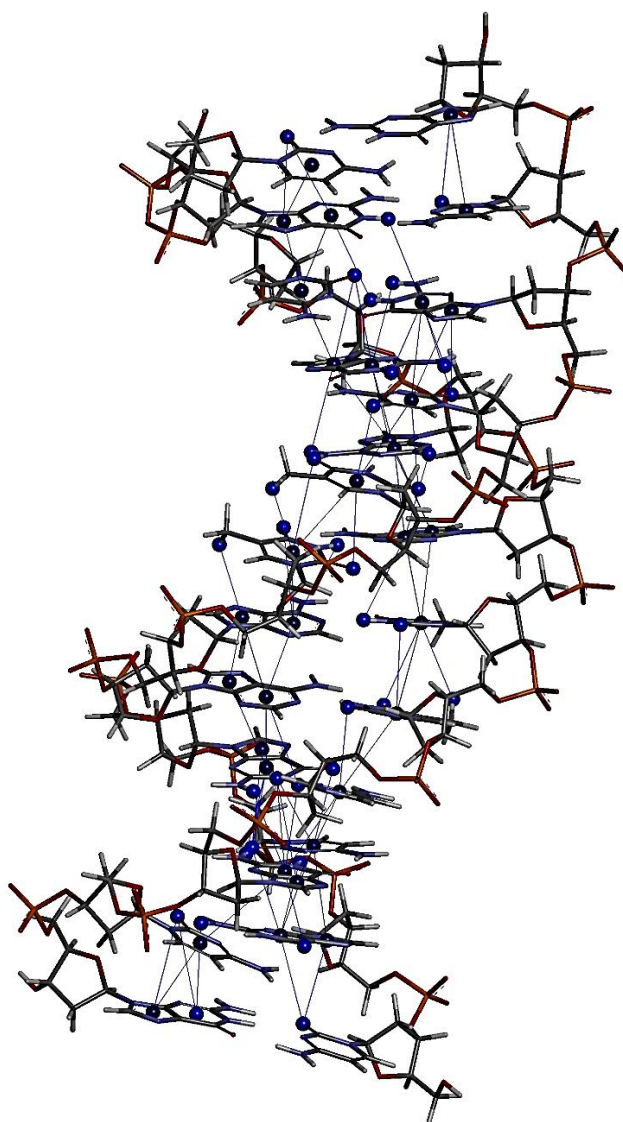
**Figura 2.5** Visualização da imagem da interação  $N\cdots\pi$  obtida por meio de cálculo vetorial de estrutura de DNA (Drew et al., 1981). Utilizado o software Discovery Studio 3.1<sup>TM</sup>, (Sacco, 2013).



**Figura 2.6** Visualização da interação  $N\cdots\pi$  obtida por meio de cálculo vetorial de estrutura de DNA (Drew et al., 1981). Utilizado o software Discovery Studio 3.1<sup>TM</sup>, (Sacco, 2013).



A Figura 2.6 mostra a acuracidade das coordenadas das interações  $\pi$  obtidas matematicamente pelo mapeamento vetorial. Tanto o centroide, como a distância e ângulos são validados pela sobreposição dos arquivos em tela gráfica, é possível visualizar a interação em seu local na estrutura cristalográfica. As imagens constantes nessa seção foram obtidas pelo processamento do arquivo de estrutura cristalográfica de DNA (Drew *et al.*, 1981), sem ligantes, com 12 pares de bases e sequência de nucleotídeos (\*CP\*GP\*CP\*GP\*AP\*AP\*TP\*TP\*CP\*GP\*CP\*G), a Figura 2.7 mostra esta estrutura com a totalidade das suas interações  $\pi$  mapeadas.



**Figura 2.7** Visualização da estrutura de DNA (Drew *et al.*, 1981) com todas as interações localizadas pelo mapeamento vetorial, utilizado o software Discovery Studio 3.1™, elaboração própria.

O procedimento de identificação das interações válidas cria tantos arquivos quanto forem os centroides existentes na estrutura analisada e podem ser estudados individualmente ou em conjunto, de forma que todas as interações fiquem sobrepostas à macromolécula e assim substituindo o trabalho manual por uma rotina computacional caracterizada por sua exatidão.

## 2.1 PLANILHAS ELETRÔNICAS – OS PRIMEIROS RESULTADOS

Ao desenvolver a pesquisa Mapeamento vetorial de interações eletrostáticas para biomoléculas, Sacco (2013) possuía como objetivo a utilização do cálculo vetorial como ferramenta para a localização dos principais tipos de interações  $\pi$ , conforme relação constante na Tabela 1.2. Assim sendo, optou-se pela seleção de estruturas de DNA, por estas apresentarem grande quantidade de anéis aromáticos. Em função disso, as estruturas tridimensionais utilizadas no trabalho de pesquisa apresentam a mesma sequência de nucleotídeos (\*CP\*GP\*CP\*GP\*AP\*AP\*TP\*TP\*CP\*GP\*CP\*G) e para efeito de estudo foram denominadas da seguinte maneira:

- DNA – SL : DNA sem ligante na estrutura (Drew et al., 1981).<sup>1</sup>
- DNA – LS : DNA com ligante no sulco da estrutura (Mallena et al., 2004).
- DNA – LI : DNA com ligante intercalado à estrutura (Malinina et al., 2002).

Todavia, ao serem analisados os primeiros resultados, vislumbraram-se aplicações anteriormente não imaginadas e relacionadas a seguir:

- Análises quantitativa e qualitativa de interações  $\pi$  em estruturas de DNA;
- Deformação da estrutura do DNA causada pela presença de ligantes;

---

<sup>2</sup> Para efeito didático, foi utilizado como referência padrão para o DNA.

- Visualização dos caminhos eletrônicos formados pelas interações  $\pi$ ;
- Variação das interações  $\pi$  em função da presença e tipo de ligantes;
- Busca por padrões de interações  $\pi$  em estruturas de DNA.

Ao analisar os dados obtidos e aplicados aos itens acima, os primeiros resultados lançaram uma nova perspectiva de estudo para a Química Supramolecular (Lehn, 1995) e DNA, descrita a seguir.

### **2.1.1 ANÁLISES QUANTITATIVA E QUALITATIVA DE INTERAÇÕES $\pi$ EM ESTRUTURAS DE DNA**

Segundo a literatura, as interações  $\pi$  são responsáveis pela conformação do DNA (Janiak, 2000; Wheeler et al., 2010a; b) e a presença de ligantes afeta as propriedades físico-químicas dos ácidos desoxirribonucleicos (Berman e Young, 1981; Edwards et al., 1992; Clausen-Schaumann et al., 2000; Williams e Maher Iii, 2000; Adams et al., 2005; Fujii et al., 2007; Nechipurenko et al., 2008), assim sendo, conclui-se que os ligantes possuem a capacidade de alterar qualitativa e quantitativamente as interações  $\pi$  existentes nas estruturas do DNA. Utilizando a metodologia desenvolvida, foram identificadas as interações  $\pi$  presentes nas estruturas de DNA sem ligante (Drew et al., 1981), na estrutura com ligante no sulco (Mallena et al., 2004) e por fim na estrutura com ligante intercalado (Malinina et al., 2002). As interações  $\pi$  totais e válidas, resultantes do mapeamento vetorial foram organizadas na Tabela 2.1 e permitem comprovar que a presença de ligantes efetua alterações qualitativas e quantitativas nas interações  $\pi$ .

**Tabela 2.1** Interações  $\pi$  identificadas em cada estrutura de DNA mapeada, incluídas as interações presentes nos ligantes. Valores absolutos das quantidades por tipo de interação, extraído de Sacco (2013)

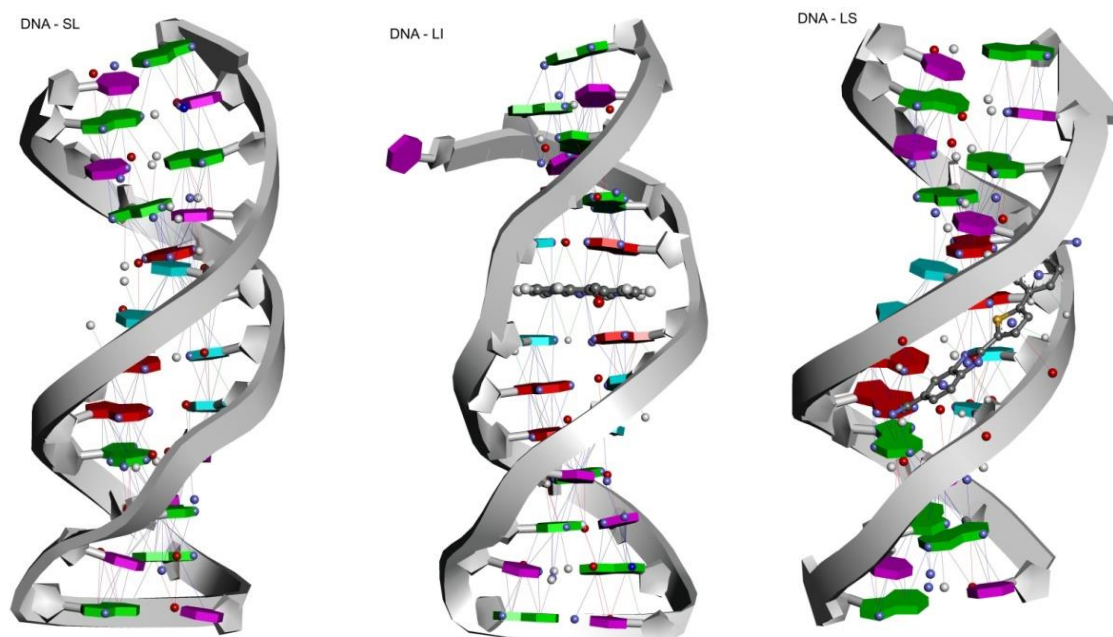
Tipo de interação $\pi$	DNA – SL	DNA – LI	DNA – LS
$N \cdots \pi$	88	83	82
$O \cdots \pi$	25	23	32
$\pi \cdots \pi$	32	33	31
$C - H \cdots \pi$	5	9	17
$N - H \cdots \pi$	14	13	13
Totais de interações por estrutura	164	161	175

É possível também expressar o resultado mostrado na Tabela 2.1 de forma relativa, reportando a proporção de cada tipo de interação em relação ao total por estrutura, conforme a Tabela 2.2.

**Tabela 2.2** Interações  $\pi$  identificadas em cada estrutura de DNA mapeada, incluídas as interações presentes nos ligantes. Valores absolutos das quantidades por tipo de interação, extraído de Sacco (2013)

Tipo de interação $\pi$	DNA – SL	DNA – LI	DNA – LS
$N \cdots \pi$	54%	52%	47%
$O \cdots \pi$	15%	14%	18%
$\pi \cdots \pi$	20%	20%	18%
$C - H \cdots \pi$	3%	6%	10%
$N - H \cdots \pi$	9%	8%	7%
Totais de interações por estrutura	100%	100%	100%

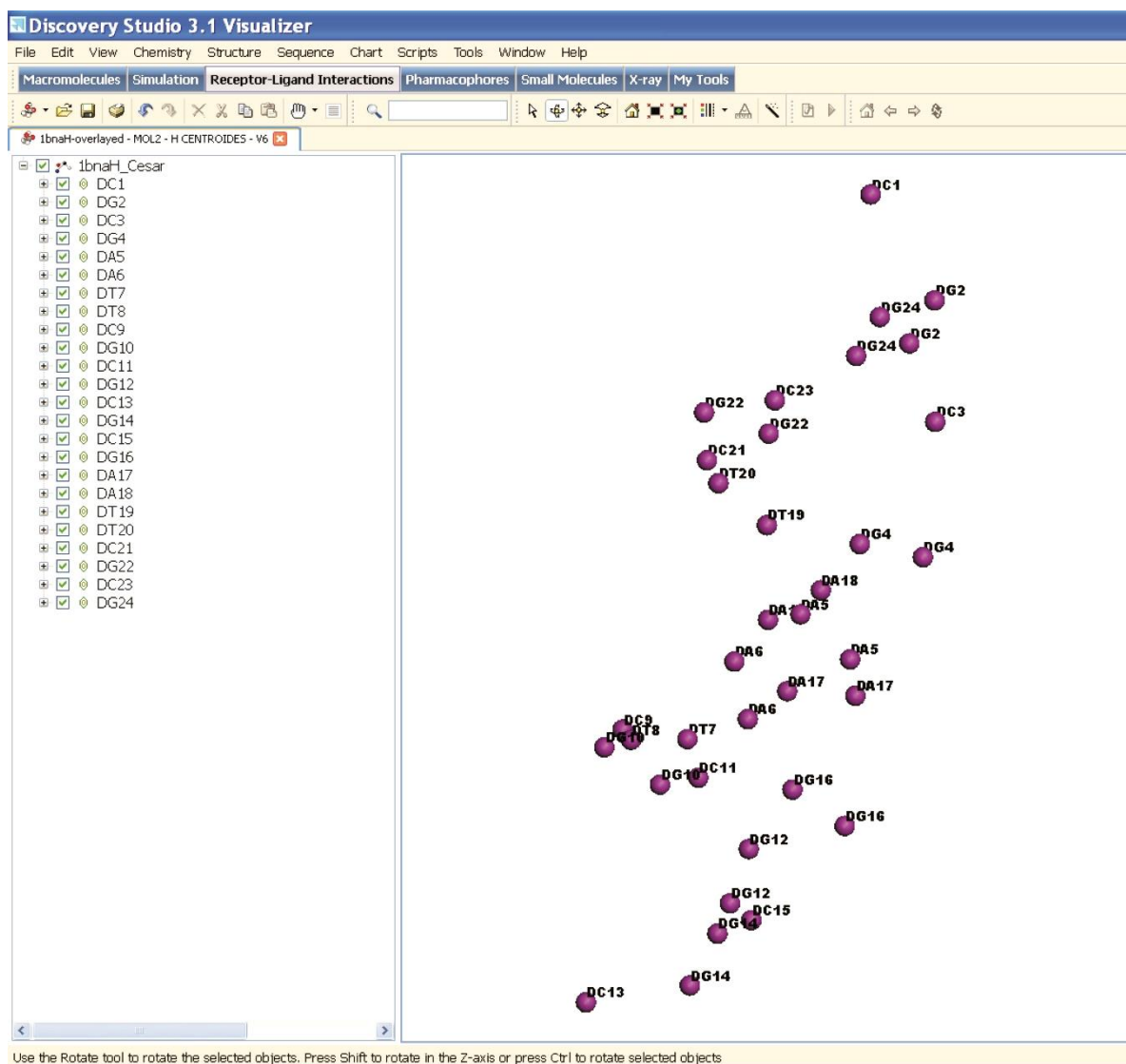
Sobrepondo as interações localizadas sobre as suas estruturas de origem, de forma análoga à mostrada na Figura 2.7, com o auxílio de software gráfico, constata-se visualmente a deformação estrutural e as alterações qualitativas e quantitativas ocorridas em função da presença dos ligantes nas estruturas do DNA, essa comparação pode ser vista na Figura 2.8.



**Figura 2.8** Imagens das estruturas de DNA com as interações  $\pi$  identificadas por mapeamento vetorial, extraído de Sacco (2013).

### 2.1.2 DEFORMAÇÃO DA ESTRUTURA DO DNA CAUSADA PELA PRESENÇA DE LIGANTES

Conforme constatado no item 2.1.2 e visualizado na Figura 2.8, a presença de ligantes altera a geometria espacial do DNA, e conseqüentemente interfere nas interações  $\pi$ . Todavia, é necessária uma melhor compreensão de onde e o quanto o ligante altera a estrutura. A resposta à esta questão surgiu ao transformar os dados de processamento intermediário mostrados na Figura 2.2 em um arquivo mol2 utilizando o procedimento da Figura 2.4. Como resultado, é obtido um arquivo compatível com os softwares para visualização em tela de computadores e que veio a tornar-se a chave para a mensuração da deformação da estrutura do DNA causada pela presença de ligantes. Trata-se da Figura 2.9, que mostra a localização espacial dos centroides de todas as bases da estrutura adotada como referência, o DNA – SL. Ao repetir o mesmo procedimento para o DNA – LI e o DNA – LS, analogamente, são obtidos outros dois arquivos com os centroides das bases pertencentes a estas estruturas cristalográficas.



**Figura 2.9** Visualização dos centroides da estrutura DNA – SL (Drew et al., 1981) obtidos por cálculo vetorial utilizando planilha eletrônica do Microsoft Excel. Visualização efetuada com a utilização do Discovery Studio 3.1™, extraído de Sacco (2013).

Tendo em vista que os arquivos no formato PDB ou similares são estruturados a partir das coordenadas  $x$ ,  $y$  e  $z$ , a deformação é mensurada em função da distância que o centroide de um mesmo anel é deslocado quando existe a presença de um ligante e utilizando a Equação 1.3 para efetuar o cálculo desse deslocamento:

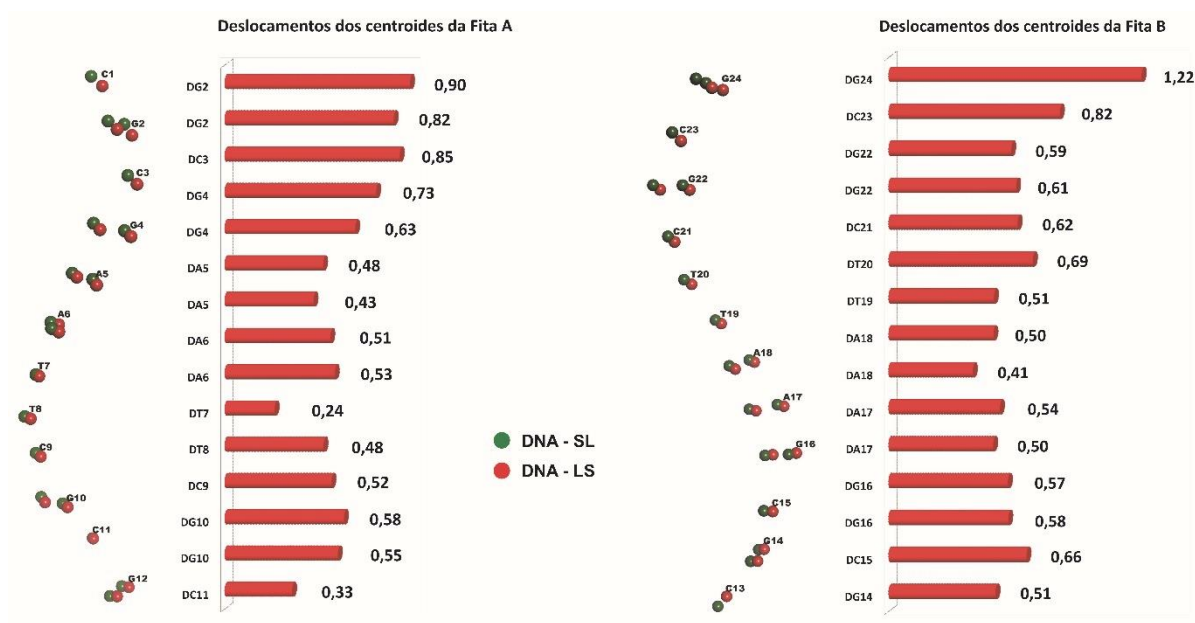
$$\text{deslocamento} = \|\overrightarrow{AB}\| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \quad (1.3)$$

Como resultado foi construída a Tabela 2.3, a qual possui os valores dos deslocamentos dos centroides das bases da estrutura de DNA com o ligante no sulco, tendo como referência a estrutura sem ligante (DNA – SL). Para melhor entendimento, a estrutura helicoidal foi dividida em Fita A e Fita B.

**Tabela 2.3** Valor do deslocamento dos centroides da estrutura DNA – LS, tendo como referência a posição do mesmo centroide na estrutura DNA – SL. O valor do deslocamento **d** é dado em Å, extraído de Sacco (2013)

Centroide	Base	d	Centroide	Base	d
1	DC1	1,00	19	DC13	1,26
2	DG2	0,90	20	DG14	0,61
3	DG2	0,82	21	DG14	0,51
4	DC3	0,85	22	DC15	0,66
5	DG4	0,73	23	DG16	0,58
6	DG4	0,63	24	DG16	0,57
7	DA5	0,48	25	DA17	0,50
8	DA5	0,43	26	DA17	0,54
9	DA6	0,51	27	DA18	0,41
10	DA6	0,53	28	DA18	0,50
11	DT7	0,24	29	DT19	0,51
12	DT8	0,48	30	DT20	0,69
13	DC9	0,52	31	DC21	0,62
14	DG10	0,58	32	DG22	0,61
15	DG10	0,55	33	DG22	0,59
16	DC11	0,33	34	DC23	0,82
17	DG12	0,57	35	DG24	1,22
18	DG12	0,59	36	DG24	1,25
Média de <b>d</b> para a fita A		0,57	Média de <b>d</b> para a fita B		0,58
Desvio padrão para <b>d</b> na fita A		0,14	Desvio padrão para <b>d</b> na fita B		0,07
Valor médio de <b>d</b> para toda a estrutura					0,58
Desvio padrão de <b>d</b> para toda a estrutura					0,11

Ao serem transformados os valores obtidos dos deslocamentos dos centroides em arquivos de texto e utilizá-los para a visualização dos deslocamentos dos centroides de ambas as estruturas de DNA em tela gráfica, nota-se de forma evidente que a deformação imposta pelo ligante no sulco não é simétrica a ambas as fitas, o que é observado não apenas pelos valores do deslocamento, mas também pelas curvas do gráfico (Figura 2.10).



**Figura 2.10** Visualização gráfica da deformação relativa calculada em função da presença de ligante no sulco da estrutura DNA – LS, tendo como referência a estrutura DNA – SL, extraído de Sacco (2013).

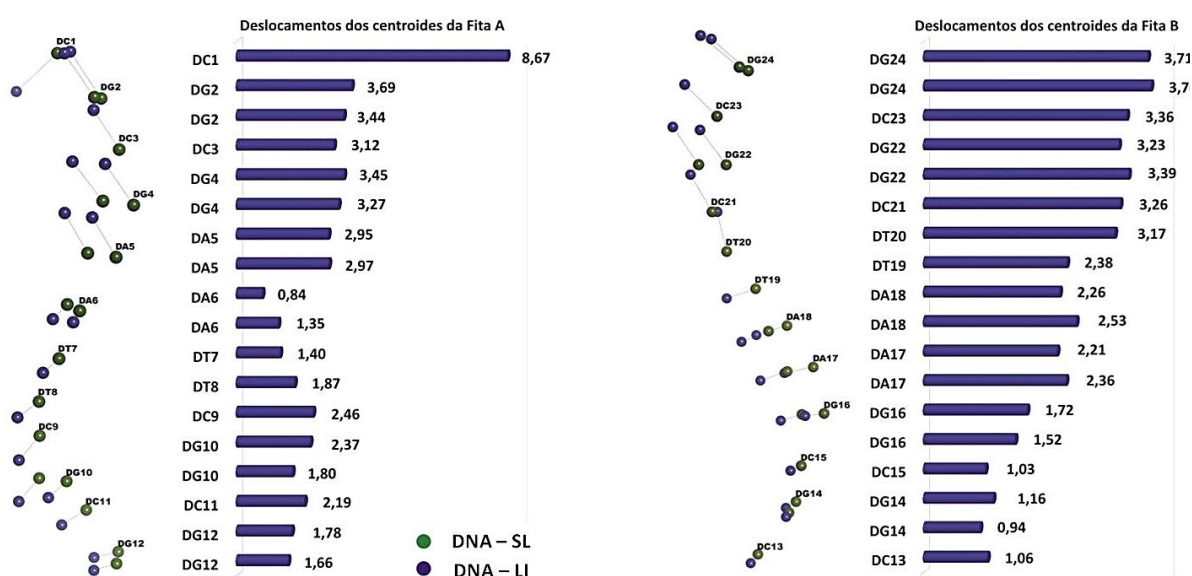
Repetindo o raciocínio anterior, são calculados da mesma forma os deslocamentos dos centroides das bases da estrutura de DNA com o ligante intercalado, tendo como referência a mesma estrutura sem ligante, DNA – SL. Os resultados estão na Tabela 2.4 e a visualização na Figura 2.11. A possibilidade de mensurar a deformação da sua estrutura, permite entender a severidade com que cada ligante age na dupla hélice do DNA.

Ao serem comparados os valores da Tabela 2.3 com os da Tabela 2.4, evidencia-se que os valores dos deslocamentos  $d$  da Tabela 2.4 além de serem maiores, apresentam uma média majorada em 4,9 vezes e o desvio padrão dos deslocamentos indica uma deformação estrutural 5,2 vezes maior.



**Tabela 2.4** Valor do deslocamento dos centroides da estrutura DNA – LI, tendo como referência a posição do mesmo centroide na estrutura DNA – SL. O valor do deslocamento **d** é dado em Å, extraído de Sacco (2013)

Centroide	Base	d	Centroide	Base	d
1	DC1	8,67	19	DC13	1,06
2	DG2	3,69	20	DG14	0,94
3	DG2	3,44	21	DG14	1,16
4	DC3	3,12	22	DC15	1,03
5	DG4	3,45	23	DG16	1,52
6	DG4	3,27	24	DG16	1,72
7	DA5	2,95	25	DA17	2,36
8	DA5	2,97	26	DA17	2,21
9	DA6	0,84	27	DA18	2,53
10	DA6	1,35	28	DA18	2,26
11	DT7	1,40	29	DT19	2,38
12	DT8	1,87	30	DT20	3,17
13	DC9	2,46	31	DC21	3,26
14	DG10	2,37	32	DG22	3,39
15	DG10	1,80	33	DG22	3,23
16	DC11	2,19	34	DC23	3,36
17	DG12	1,78	35	DG24	3,76
18	DG12	1,66	36	DG24	3,71
Média de <b>d</b> para a fita A		2,78	Média de <b>d</b> para a fita B		2,30
Desvio padrão para <b>d</b> na fita A		0,74	Desvio padrão para <b>d</b> na fita B		0,70
Valor médio de <b>d</b> para toda a estrutura					2,39
Desvio padrão de <b>d</b> para toda a estrutura					0,72

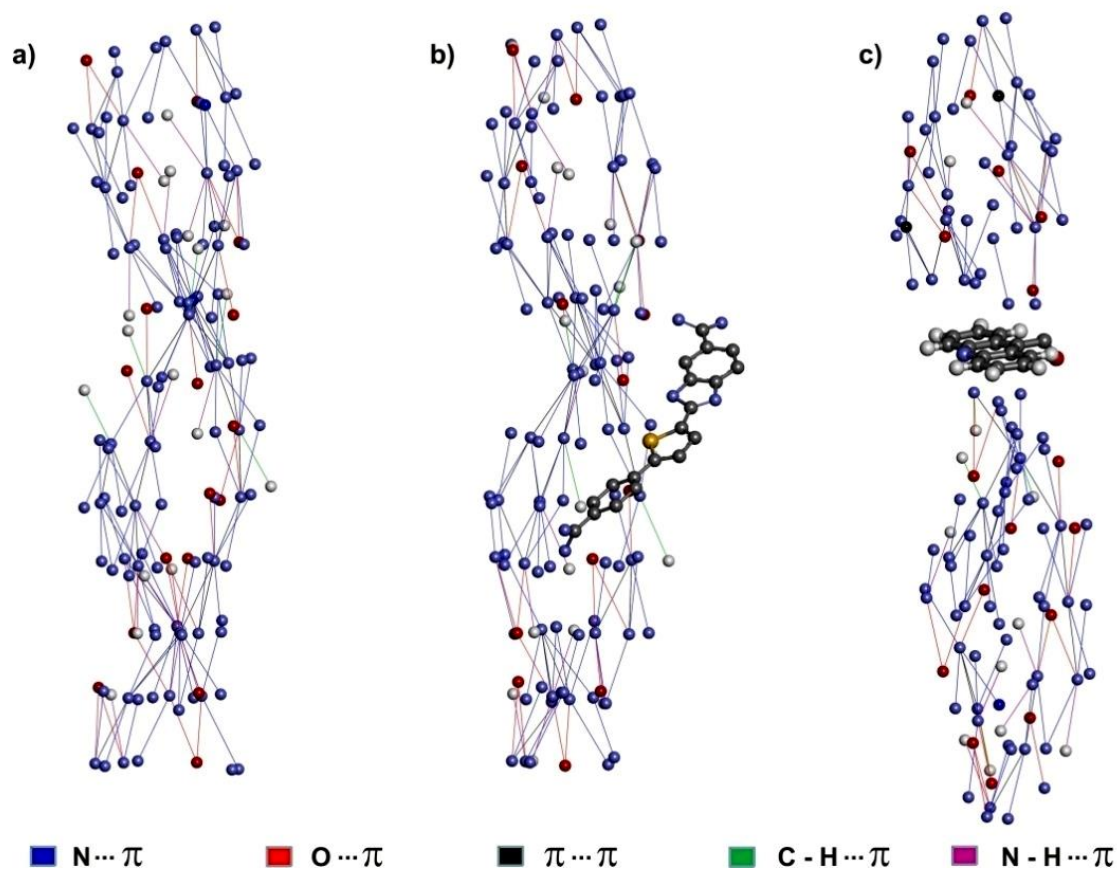


**Figura 2.11** Visualização gráfica da deformação relativa calculada em função da presença de ligante intercalado na estrutura DNA – LI, tendo como referência a estrutura DNA – SL, extraído de Sacco (2013).

### 2.1.3 VISUALIZAÇÃO DOS CAMINHOS ELETRÔNICOS FORMADOS PELAS INTERAÇÕES $\pi$

Conforme mostradas anteriormente na Figura 2.7 todas as interações  $\pi$  determinadas para o DNA – SL (Drew et al., 1981) sobrepostas à sua estrutura original pode-se observar que o conjunto de todas as interações  $\pi$  ( $N\cdots\pi$ ,  $O\cdots\pi$ ,  $C-H\cdots\pi$ ,  $N-H\cdots\pi$ ,  $O-H\cdots\pi$  e  $\pi\cdots\pi$ ) forma uma rede de caminhos eletrostáticos. Sabendo-se que as interações  $\pi$  contribuem para a conformação e estabilidade das proteínas (Wintjens et al., 2000; Cauët et al., 2005) e ainda segundo alguns pesquisadores (Hunter e Sanders, 1990; Hunter et al., 2001; Kiviniemi, 2001), esse fenômeno ocorre por meio da relação doação e aceitação de cargas eletrostáticas. Com base nessas premissas surgiu a ideia de visualizar em tela gráfica todas as interações obtidas para as demais estruturas de DNA (DNA – LS e DNA – LI), todavia para melhor observação destes caminhos, optou-se por não inserir nas imagens as estruturas cristalográficas das três estruturas. A Figura 2.12 mostra as relações de transferência de cargas eletrostáticas em todos os sistemas  $\pi$  e demonstra como a presença de ligantes interfere nestes, alterando a sua configuração. As interações entre os ligantes e estruturas de DNA foram suprimidas.

As imagens obtidas, vêm de encontro ao trabalho de diversos pesquisadores que definem as interações eletrostáticas como meio de fixação dos ligantes às estruturas do DNA. A literatura pesquisada definiu os ligantes não apenas como fármacos, mas também toxinas e duas principais formas de ligação. Na primeira destas, alguns fármacos fixam-se por intercalação na dupla hélice (Gao et al., 1991; Cerny e Hobza, 2007; Shaikh e Jayaram, 2012) e também o fazem dessa forma algumas toxinas (Ferguson e Denny, 2007) e a outra forma é a acomodação do ligante no sulco menor (Wellenzohn et al., 2001; Nelson et al., 2007; Shaikh e Jayaram, 2012). As imagens a seguir, mostram a alteração dos caminhos eletrostáticos em função dos ligantes e conseqüentemente a funcionalidade das estruturas estudadas, uma estrutura sem ligantes foi adicionada ao quadro para servir como referência.



**Figura 2.12** Diagrama de transferência de cargas eletrostáticas nos sistemas  $\pi$  nas estruturas de DNA. a) DNA – SL; b) DNA – LS e c) DNA – LI. Utilizado o software Discovery Studio 3.1™ para visualização.

#### 2.1.4 VARIAÇÃO DAS INTERAÇÕES $\pi$ EM FUNÇÃO DA PRESENÇA E TIPO DE LIGANTES

Ainda como resultado obtido em pesquisa anterior, os dados obtidos pelo mapeamento vetorial foram disponibilizados em dois tipos de arquivos conforme dito anteriormente, são arquivos compatíveis com planilhas eletrônicas e arquivos para utilização em *graph sets*, (Sacco, 2013). Considerando-se que todos os arquivos que geraram as figuras mostradas anteriormente e tabelas foram também disponibilizados na extensão **.xlsx** para o Microsoft Excel (Microsoft.Com, 2012a) e conhecendo-se as ferramentas de análise de dados disponíveis nesse *software*, foi identificada uma nova oportunidade de estudo das interações  $\pi$  (N... $\pi$ , O... $\pi$ , C-H... $\pi$ , N-H... $\pi$ , O-H... $\pi$  e  $\pi$ ... $\pi$ ). Embora a quantidade total das interações por estrutura (DNA – SL, DNA – LS e DNA – LI)

mostradas na Tabela 2.1 não possui uma variação expressiva, o tipo de posicionamento do ligante apresenta uma significativa alteração no tipo de interação existente em função da conformação do DNA em função dos ligantes. Utilizando os dados obtidos e com os recursos da tabela dinâmica existentes na planilha eletrônica Excel, foi possível identificar os totais de interações  $\pi$  mapeadas para as três estruturas cristalográficas e identificar o tipo de interação por meio dos átomos ou centroides relacionados a cada base e também a sua posição nesta, onde podem ser observadas as variações por meio dos valores do comprimento da interação e ângulo  $\alpha$ . Dessa forma, pode-se quantificar o quanto a presença do ligante e a forma como este se posiciona na estrutura, afeta as propriedades físico-químicas destas (Nechipurenko et al., 2008). Como a intenção deste tópico é demonstrar os primeiros resultados obtidos com o Mapeamento vetorial de interações eletrostáticas para biomoléculas, serão reproduzidas apenas duas tabelas extraídas do trabalho em tela, Tabelas 2.5 e 2.6, pois foram consideradas apenas as interações  $\pi$  com origem nos centroides do DNA.

**Tabela 2.5** Interações  $\pi$  com origem na base G2, anel aromático com 6 átomos, com os átomos e centroides relacionados por bases com o respectivo comprimento  $d$  da interação e ângulo  $\alpha$ , extraído de Sacco (2013)

Centroide de origem da interação: G2-A								
Interage com	Átomo	Tipo de interação	DNA – LI		DNA – LS		DNA – SL	
			$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)
C1	O2	O $\cdots$ $\pi$			3,71	7,06	3,63	1,84
	N1	N $\cdots$ $\pi$			4,02	28,66		
C3	N3	N $\cdots$ $\pi$			4,01	28,78		
	N4	N $\cdots$ $\pi$	4,12	19,87	3,97	16,03	3,89	21,98
	CT6	$\pi$ $\cdots$ $\pi$			3,86	26,66	3,94	27,45
G24	N2	N $\cdots$ $\pi$	3,55	28,04				

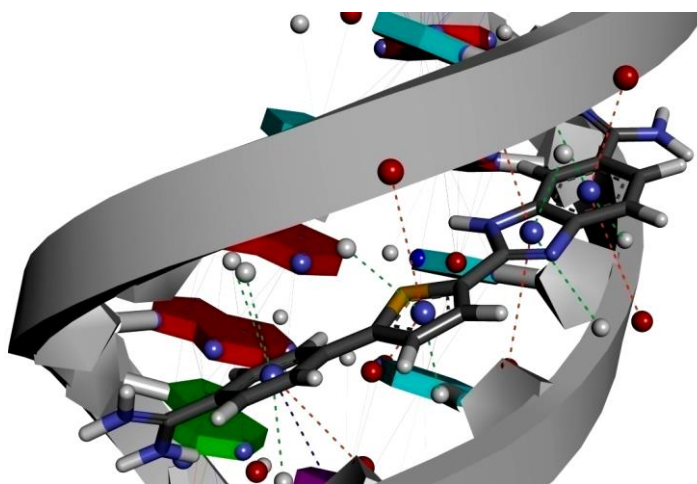
A tabela anterior apresenta uma interação N $\cdots$  $\pi$  entre o centroide do anel aromático com seis átomos da guanina G2 com o nitrogênio N4 da citosina C3 comum às três estruturas de DNA, todavia, em função da presença dos ligantes há alteração do comprimento desta interação e do valor do ângulo existente entre o vetor normal ao

plano do centroide e o vetor distância centroide-átomo. A seguir, a Tabela 2.6, que contém interações com o ligante intercalado.

**Tabela 2.6** Interações  $\pi$  com origem na base T20, com os átomos e centroides relacionados por bases com o respectivo comprimento  $d$  da interação e ângulo  $\alpha$ , extraído de Sacco (2013)

Centroide de origem da interação: T20								
Interage com	Átomo	Tipo de interação	DNA – LI		DNA – LS		DNA – SL	
			$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)
9AC700	CT6	$\pi \cdots \pi$	3,55	16,27				
	H2	C - H $\cdots \pi$	3,63	29,36				
	H3	N - H $\cdots \pi$	3,68	25,87				
C21	CT6	$\pi \cdots \pi$			3,98	25,46	3,94	26,21
	N4	N $\cdots \pi$	3,42	22,90	3,76	26,57	3,66	26,93
	H5	C - H $\cdots \pi$			3,71	18,56	3,77	17,60
T19	CT6	$\pi \cdots \pi$			3,90	26,20	4,09	28,90
	N3	N $\cdots \pi$			3,67	24,67	3,82	24,67
	O2	O $\cdots \pi$			3,55	15,08	3,60	11,01
	N1	N $\cdots \pi$			3,95	23,17	4,14	27,05

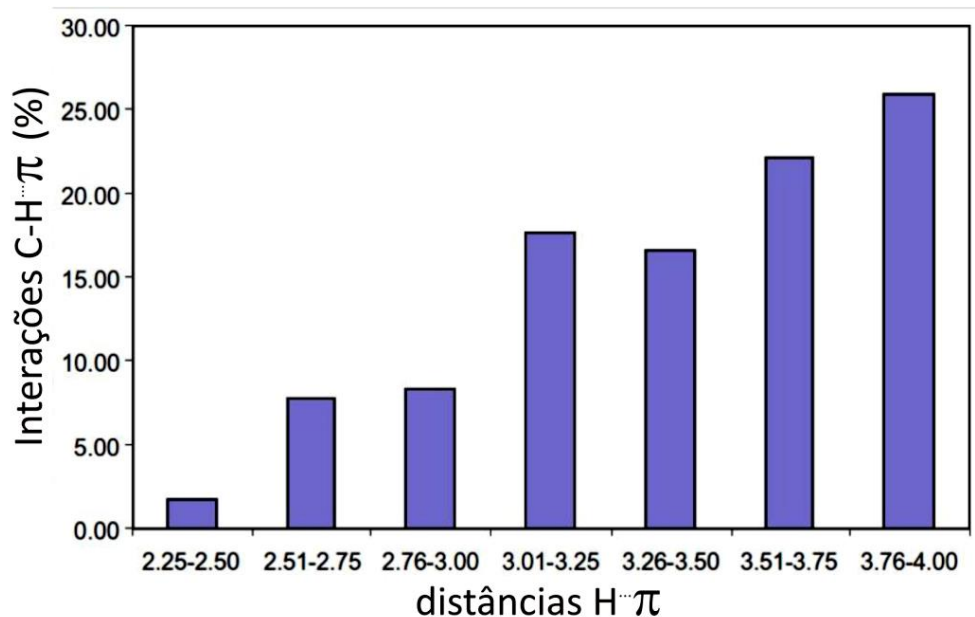
Podem ser observadas na Tabela 2.6 as interações da base T20 com o ligante 9AC700. Não foram geradas tabelas semelhantes que contenham as interações  $\pi$  com ligante no sulco, não existem interações  $\pi$  para este com origem no DNA, conforme pode ser visto na Figura 2.10.



**Figura 2.13** Representação gráfica das interações  $\pi$  entre o ligante e as bases de uma estrutura de DNA – LS (Mallena et al., 2004). Estrutura gerada e visualizada com o software Discovery Studio 3.1™, extraído de Sacco (2013).

### 2.1.5 BUSCA POR PADRÕES DE INTERAÇÕES $\pi$ EM ESTRUTURAS DE DNA

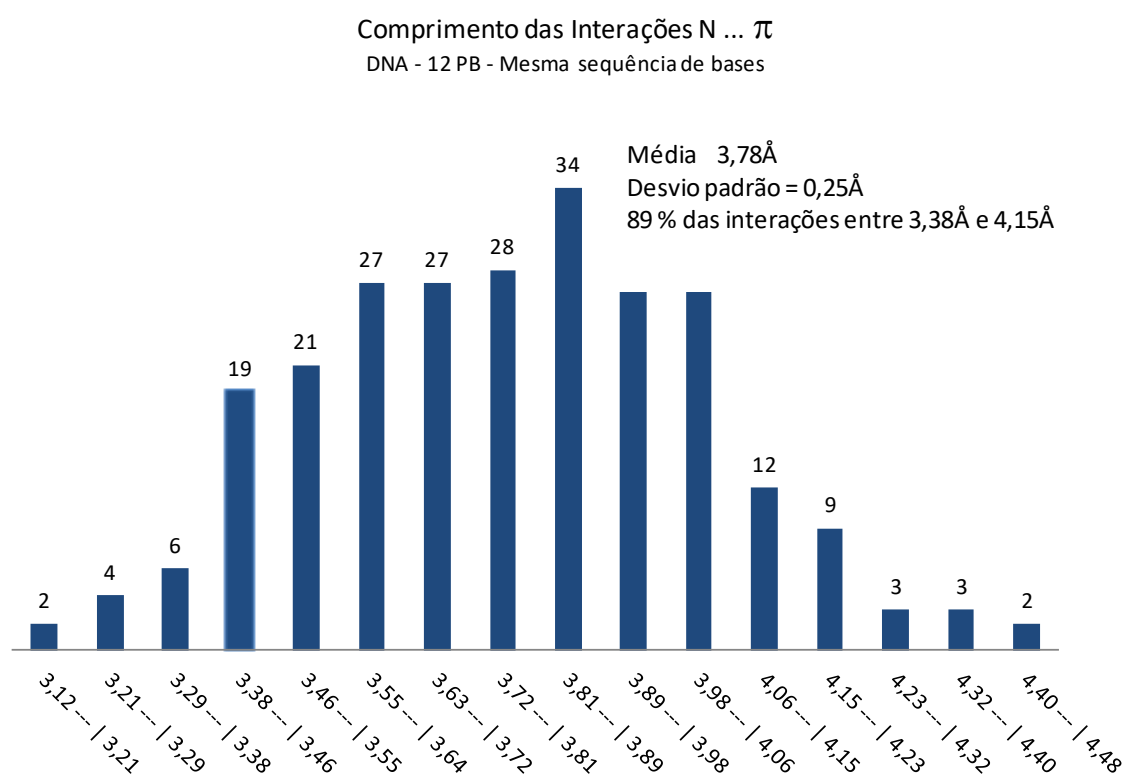
Durante a pesquisa relacionada à dissertação de mestrado, um artigo de Shanthi e colaboradores (Shanthi et al., 2010) deu origem à uma nova aplicação para o Mapeamento Vetorial para determinação das interações  $\pi$  em macromoléculas e DNA. Em seu trabalho, Shanthi efetuou uma análise estatística sobre as interações C-H $\cdots$  $\pi$  existentes na cadeia lateral de uma amostragem de 75 proteínas e a partir dos dados obtidos tabulou o comprimento das interações entre os átomos de hidrogênio e os centroides, Figura 2.14. O resultado por ele apresentado direcionou a aplicação dos dados já existentes e utilizados na análise da variação das interações  $\pi$  em função da presença e tipo de ligantes (Item 2.1.4) na busca por padrões de interações  $\pi$  em estruturas de DNA.



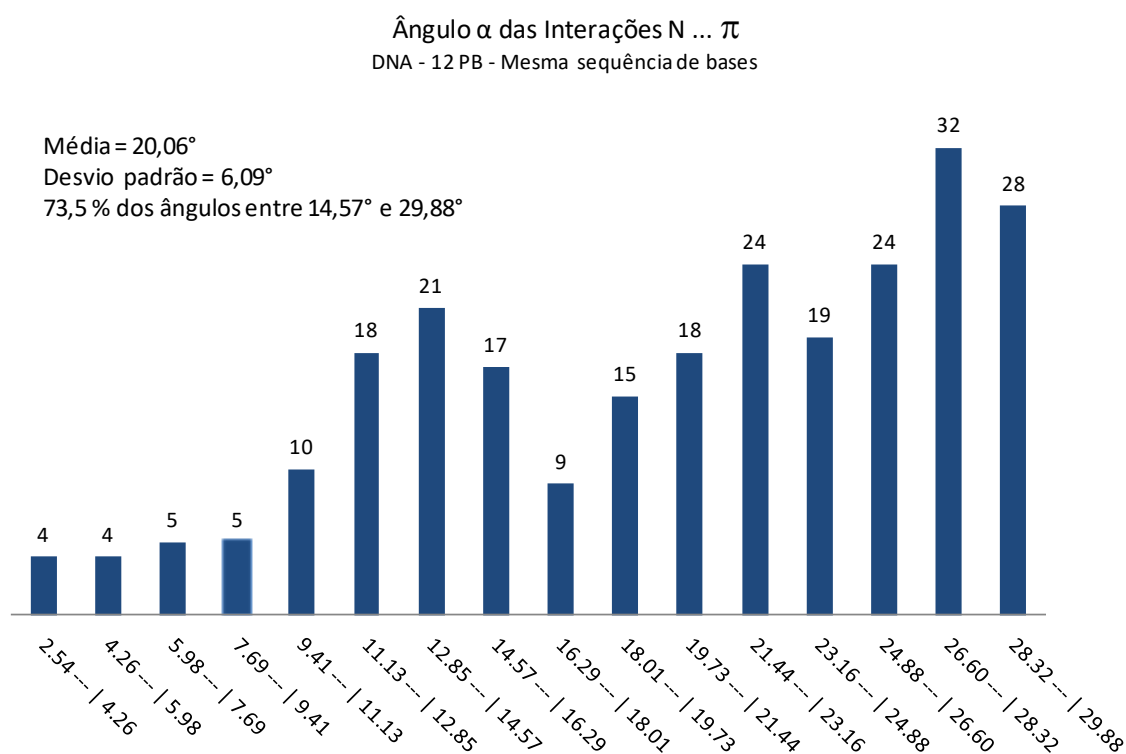
**Figura 2.14** Representação gráfica da tabulação das distâncias entre os átomos de hidrogênio e centroides nas interações C-H $\cdots$  $\pi$  da cadeia lateral de proteínas. Extraído de Shanthi e colaboradores (2010).

Como resultado, os cálculos retornam os valores dos parâmetros geométricos das distâncias e ângulos do vetor distância relativos ao vetor normal ao plano do anel, os quais quando submetidos à uma análise estatística por tipo de interação  $\pi$  (N $\cdots$  $\pi$ , O $\cdots$  $\pi$ , C-

$H\cdots\pi$ ,  $N-H\cdots\pi$ ,  $O-H\cdots\pi$  e  $\pi\cdots\pi$ ), possibilita verificar se para as estruturas estudadas existe ou não um padrão para as distâncias e ângulos. Com os recursos disponíveis do Microsoft Excel, para cada sequência de dados analisada foi possível conhecer a média dos parâmetros geométricos, o desvio padrão da amostra e o intervalo na qual estão concentradas a maioria das medidas obtidas e assim observar graficamente se há um padrão ou não para as interações localizadas. Para ilustrar mais essa aplicação do Mapeamento Vetorial, encontram-se exemplificados pelas Figuras 2.15 e 2.16, nas quais estão plotadas as curvas das frequências das distâncias e ângulos obtidos para as interações  $N\cdots\pi$  existentes nas três estruturas de DNA (DNA – SL, DNA – LS e DNA – LI) estudadas.



**Figura 2.15** Representação gráfica da tabulação das distâncias das interações  $N\cdots\pi$  das estruturas DNA – SL, DNA – LS e DNA – LI, (Sacco, 2013).



**Figura 2.16** Representação gráfica da tabulação dos ângulos entre o vetor distância e o vetor normal ao plano do anel aromático das interações N... $\pi$  das estruturas DNA – SL, DNA – LS e DNA – LI, (Sacco, 2013).

É interessante observar que para as distâncias vetorialmente determinadas envolvidas nas interações N... $\pi$ , obtém-se uma curva normal e para os ângulos dessas mesmas interações nota-se uma curva trimodal. Em sua dissertação de mestrado, Sacco (2013), efetuou a mesma análise estatística para a busca de padrões também nas demais interações  $\pi$  (O... $\pi$ , C-H... $\pi$ , N-H... $\pi$ , O-H... $\pi$  e  $\pi$ ... $\pi$ ).

## 2.2 O MAPEAMENTO VETORIAL NA PESQUISA E PRODUÇÃO ACADÊMICAS

Além dos resultados obtidos com o Mapeamento Vetorial anteriormente descrito neste capítulo, as planilhas eletrônicas utilizadas vieram a constituir valiosa contribuição para os colaboradores do LaCrEMM ( Zukerman-Schpector, J.; Caracelli, I.; Haiduc, I., 2012) e BioMat ( Caracelli, I.; Zukerman-Schpector, J.; Araujo-Moreira, F. M.,



2012) em pesquisas para dissertações de mestrado e tese de doutorado. Além disso, juntamente com esses dois laboratórios, foi desenvolvida uma pesquisa de Iniciação Científica conjuntamente com a Faculdade de Tecnologia de Tatuí. A Tabela 2.7 mostra uma tabulação do Mapeamento Vetorial aplicado às pesquisas citadas e também os trabalhos acadêmicos relacionados à esta ferramenta e apresentados em eventos. A relação da pesquisa e produção acadêmicas está listada em ordem cronológica, e por tipo nos itens a seguir (2.2.1, 2.2.2 e 2 2.2.3).

**Tabela 2.7** Trabalhos de pesquisa, produção acadêmica e programa de computador decorrentes do Mapeamento Vetorial, elaboração própria.

Teses de Doutorado	1
Dissertações de Mestrado	2
Iniciação Científica	1
Trabalhos apresentados em eventos	8
Programa de computador	1

### 2.2.1 PROJETOS DE PESQUISA

Leves, Natália. **Nanopartículas de grafite para carreamento de anti-inflamatórios não esteroidais por estudos de *docking* molecular.** 2013. Dissertação (Mestrado em Programa de Pós-graduação em Biotecnologia) – UFSCar, São Carlos, 2013.

Pereira, Fernanda. **Mapeamento vetorial para interações  $\pi$  em macromoléculas: busca por padrões em grupos específicos de proteínas.** 2013. Iniciação Científica. (Graduando em Gestão da Tecnologia da Informação) - Faculdade de Tecnologia de Tatuí, Tatuí, 2013.

Baptistini, Natália. **Análise *in silico*, *in vitro* e *in vivo* de compostos organocalcogênicos como possíveis anti-inflamatórios.** 2015. Tese (Doutorado em Programa de Pós-graduação em Biotecnologia) – UFSCar, São Carlos, 2015.

Silva, Rui Filipe Nogueira da. **Análise das Características Estruturais do FAD em Oxidorredutases**. 2015. Dissertação (Mestrado em Programa de Pós-graduação em Biotecnologia) – UFSCar, São Carlos, 2015.

### 2.2.2 TRABALHOS APRESENTADOS E PUBLICADOS

Sacco, A. C. S.; Caracelli, I.; Zukerman-Schpector, J.: **Mapeamento vetorial de interações eletrostáticas para macromoléculas**. 1.º Encontro Nacional da Rede Nacional de Nanobiotecnologia Aplicada à Medicina e à Defesa, 2010, São Carlos. PPG Biotec - UFSCar. p.7.

Caracelli, I.; Zukerman-Schpector, J.; Sacco, A. C. S.; Pereira, F.: **A informática na pesquisa em biomoléculas**. In: (Ed.). Tatuí: FATEC Tatuí, 2013. p.55-57. (VIII Simpósio de Ciência e Tecnologia - Caderno de Resumos).

Caracelli, I.; Zukerman-Schpector, J.; Sacco, A. C. S.; Pereira, F.: **Mapeamento vetorial de interações  $\pi$  em macromoléculas: busca por padrões em grupo específico de DNA**. In: (Ed.). Tatuí: FATEC Tatuí, 2014. p.75-76. (IX Simpósio de Ciência e Tecnologia - Caderno de Resumos).

Sacco, A. C. S.; Souza, A. L.; Caracelli, I.; Zukerman-Schpector, J.: M. **Aplicação de técnicas computacionais ao estudo de biomoléculas**. In: (Ed.). Tatuí: FATEC Tatuí, 2014. p.9-10. (IX Simpósio de Ciência e Tecnologia - Caderno de Resumos).

Caracelli, I.; Zukerman-Schpector, J.; Sacco, A. C. S.; Pereira, F.: **Mapeamento vetorial de interações  $\pi$  em macromoléculas: busca por padrões em grupo específico de DNA**. Revista Sapere. Tatuí: FATEC Tatuí. 6: 1-13 p. 2014.

Caracelli, I.; Zukerman-Schpector, J.; Sacco, A. C. S.; Souza, L. M.: **Requisitos de um projeto de software para o cálculo de interações  $\pi$ .**

In: (Ed.). Tatuí: FATEC Tatuí, 2015. p.145-146. (X Simpósio de Ciência e Tecnologia - Caderno de Resumos).

Caracelli, I.; Zukerman-Schpector, J.; Sacco, A. C. S.; Souza, L. M.: **Desenvolvimento de software para cálculo de interações  $\pi$ .** Revista Sapere. Tatuí: FATEC Tatuí. 8: 1-5 p. 2016.

Sacco, A. C. S.; Zukerman-Schpector, J.; Caracelli, I.: **Aplicação do mapeamento vetorial em DNA.** In: (Ed.). São Carlos: UFSCar - PPG Biotec - UFSCar, 2016. p.5-6. (V Simpósio de Biotecnologia da Universidade Federal de São Carlos - Caderno de Resumos).

### 2.2.3 PROGRAMA DE COMPUTADOR

Sacco, A. C. S.; Caracelli, I.; Zukerman-Schpector, J.; Souza, L. M.: **WIM - Weak Interaction Mapping.** Tatuí 2016.

## 2.3 DESENVOLVIMENTO DE UM SOFTWARE PARA O MAPEAMENTO VETORIAL

Ao relacionarem-se os resultados acima, onde o modelo matemático concebido para o mapeamento das interações  $\pi$  utiliza cálculo vetorial aplicado às coordenadas cristalográficas dos átomos de proteínas e DNA (Sacco, Zukerman-Schpector, *et al.*, 2016) e identificando nestes um evidente potencial para a transformação do Mapeamento Vetorial em uma ferramenta, decidiu-se pela sua total automação, transformando o seu algoritmo experimental (Sacco, 2013) em um programa de computador que além de mapear e caracterizar as interações  $\pi$  ( $\text{CH}\cdots\pi$ ,  $\text{NH}\cdots\pi$ ,  $\pi\cdots\pi$  e outras), também fizesse o mesmo com as ligações de hidrogênio. O sistema batizado de Weak Interaction Mapping (WIM), teve seu algoritmo aprimorado diversas vezes buscando otimizar o seu tempo de processamento e tornar seu visual intuitivo (Caracelli,

I. et al., 2015; Caracelli, I. et al., 2016; Sacco, A. C. S.; et al., 2016) e após 18 meses em desenvolvimento e testes, teve o pedido de propriedade intelectual para programa de computador depositado na Agência de Inovação (FAI, 2017) da UFSCar sob o processo número PC 2017/004. A seguir, serão comentadas resumidamente as etapas de desenvolvimento do *software*.

### **2.3.1 REQUISITOS BÁSICOS DO SOFTWARE**

A definição dos requisitos de um software ou sistema está relacionada com a identificação dos problemas existentes a serem resolvidos (Laudon e Laudon, 2004), neste caso, o *software* deve ser capaz de selecionar uma estrutura cristalográfica de uma biomolécula (DNA ou proteína), e utilizando o mapeamento vetorial, determinar a localização e os tipos de interações  $\pi$  e/ou ligações de hidrogênio existentes e como produto final retornar arquivos destas, compatíveis com planilhas eletrônicas e/ou visualizadores gráficos (Caracelli, I. et al., 2015; Caracelli, I. et al., 2016).

### **2.3.2 DEFINIÇÃO DA LINGUAGEM DE PROGRAMAÇÃO**

Considerando a evolução dos recursos computacionais, tornou-se necessária uma alternativa que substituísse a programação em linguagem de máquina, e como solução para esse problema surgiram as linguagens de programação, as quais tornam possível escrever programas utilizando instruções mais compreensíveis ao ser humano (BROOKSHEAR, 2013). Para o desenvolvimento do WIM (Caracelli, I.; Zukerman-Schpector, J.; Souza, L. M.) foi utilizada a linguagem de programação C#, que é uma linguagem de programação para desenvolvimento em .NET Framework (Microsoft, 2010). Embora essa LP (linguagem de programação) possua algumas características típicas de linguagens de programação imperativas, C# é uma linguagem de programação orientada a objeto, pois ela suporta herança, polimorfismo e dados de tipo abstrato (Sebesta, 2012). Para o desenvolvimento de softwares existem diversos tipos de ambientes de programação, tais como sistemas de arquivo, editores de texto e compiladores (Sebesta,

2012) e para tornar o desenvolvimento desse aplicativo mais produtivo, foi utilizado um ambiente integrado de desenvolvimento (IDE). Para atender às especificações do projeto, a IDE utilizada foi o Microsoft Visual Studio 2010 (Microsoft, 2010), pois essa IDE possui uma estrutura robusta e proporciona praticidade ao desenvolvimento na linguagem de programação C# (Sharp, 2010; Sebesta, 2012).

### 2.3.3 INTERFACE COM O USUÁRIO

O WIM é um sistema de uso altamente intuitivo e possui uma tela inicial simples (Figura 2.17), onde como os *inputs* para o processamento são:

- Tipo de arquivo a processar (pdb ou mol2);
- Localização do arquivo, no website do PDB (Rcsb, 2012) ou em unidade física de armazenamento;
- Tipo de arquivo a processar (pdb ou mol2);
- Localizar as ligações de hidrogênio;
- Localizar as interações  $\pi$ ;
- Localizar as ligações de hidrogênio e as interações  $\pi$ .

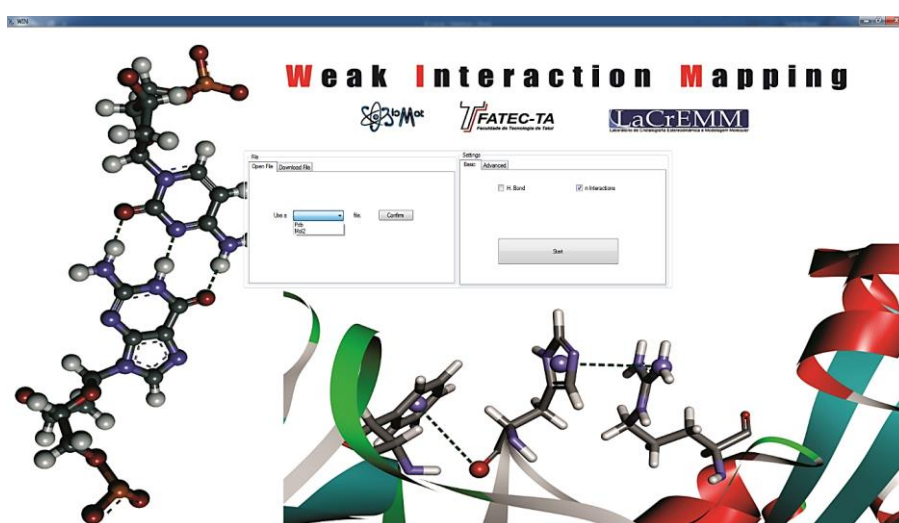
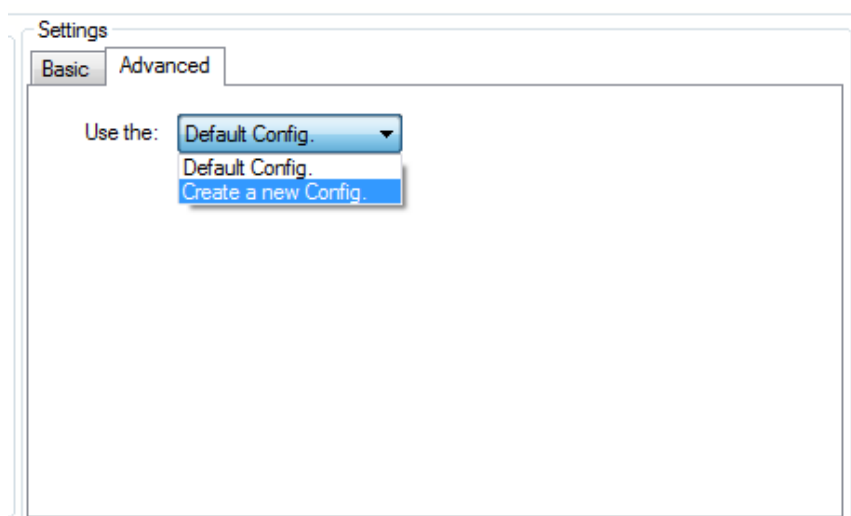
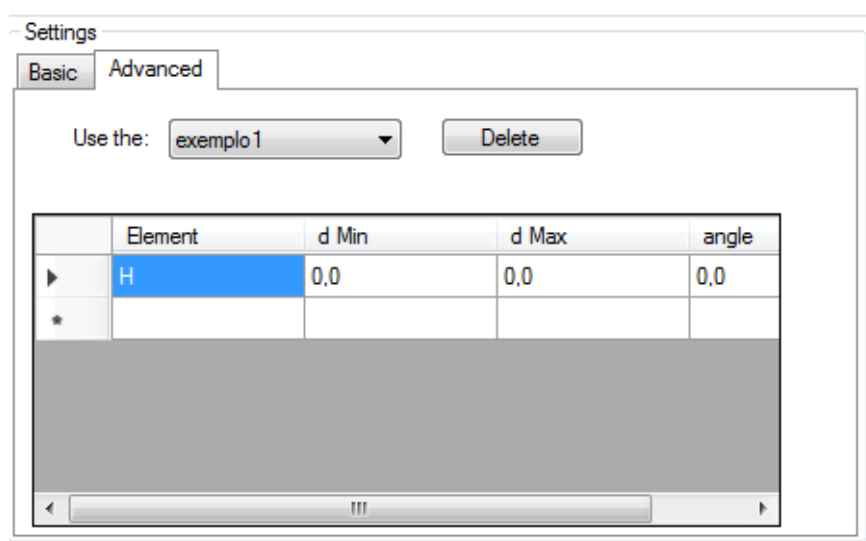


Figura 2.17 Tela inicial do WIM (Caracelli, I. et al., 2016).

Dessa forma, o software executará as necessidades específicas do usuário. Embora os parâmetros geométricos considerados como default para os cálculos necessários para o mapeamento das interações não covalentes sejam os valores das Tabelas 1.2 e 1.4, há a possibilidade de personalização dos parâmetros para interações específicas utilizando as opções avançadas no painel de interações, conforme exemplificado pelas Figuras 2.18 e 2.19.



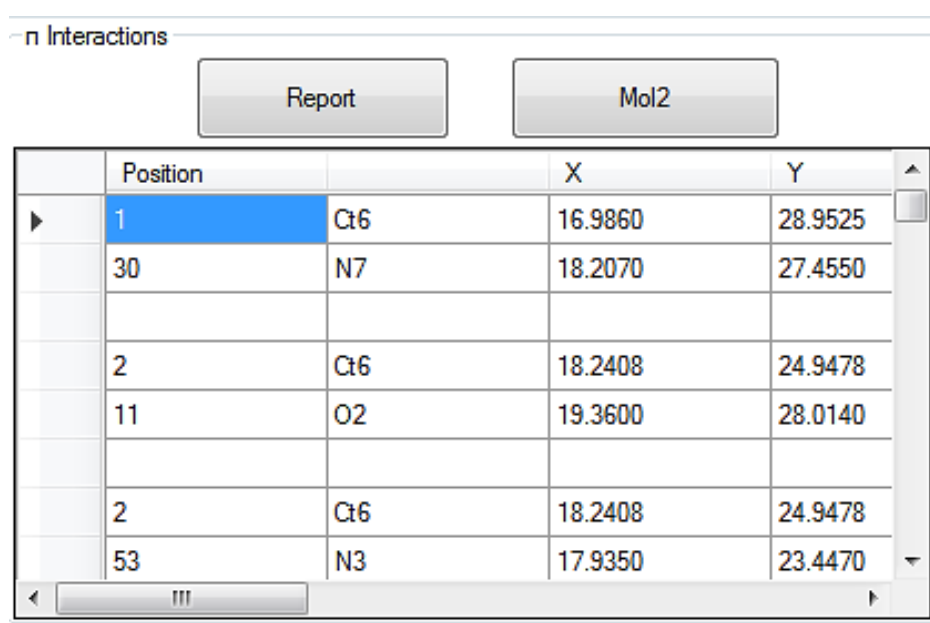
**Figura 2.18** Tela de criação de configuração para o cálculo das interações  $\pi$ , WIM (Caracelli, I. et al., 2016).



**Figura 2.19** Tela de configuração ajustada para os parâmetros personalizados, WIM (Caracelli, I. et al., 2016).

### 2.3.4 OUTPUT DO PROCESSAMENTO

Como output do processamento, há uma tela de interface com o usuário que permite além da exportação do resultado em arquivos mol2, em arquivo para planilha eletrônica, ambas as opções anteriores ou simplesmente a visualização das interações em tela. A Figura 2.20 utiliza como exemplo as interações  $\pi$  identificadas em uma estrutura do DNA.



π Interactions

Report Mol2

	Position		X	Y
▶	1	Ct6	16.9860	28.9525
	30	N7	18.2070	27.4550
	2	Ct6	18.2408	24.9478
	11	O2	19.3600	28.0140
	2	Ct6	18.2408	24.9478
	53	N3	17.9350	23.4470

**Figura 2.20** Ilustração da visualização das interações  $\pi$  resultantes do processamento pelo WIM (Caracelli, I. et al., 2016)

A criação do WIM, além de automatizar o trabalho de identificação das interações  $\pi$  e Ligações de hidrogênio com precisão, reduz o tempo despendido na identificação destas interações e conforme exposto neste capítulo, permitirá o acesso a informações que até o presente momento não eram disponíveis e dessa forma, constituindo uma ferramenta que vem a preencher uma lacuna existente até então no campo da pesquisa de uma das interações eletrostáticas importantes em macromoléculas.

### 2.3.5 CONSIDERAÇÕES SOBRE O DESENVOLVIMENTO

Para a obtenção do algoritmo final do WIM foram necessárias que algumas questões-chave viessem a ser solucionadas, sendo que estas serão abordadas conforme a sequência lógica do algoritmo.

#### 2.3.5.1 CONVERSÃO DE ARQUIVOS .PDB EM .MOL2

De acordo com o disposto no Item 2.3.3, o WIM aceita como *input* tanto arquivos com extensão .pdb como .mol2. Embora ambas as extensões sejam variações de arquivos de texto, o reconhecimento destes por arquivos de visualização gráfica é efetuado pelo formato *layout* dos dados, possuindo a extensão .mol2 um formato mais simples, semelhante à uma tabela com um cabeçalho contidos em cada arquivo conforme mostrado na Tabela 1.5 e Figura 2.4. Objetivando redução do tempo de processamento padrão, optou-se que as rotinas internas processarem arquivos .mol2. Para tal, foi necessário o desenvolvimento de um algoritmo responsável pela conversão de arquivos no formato .pdb em .mol2.

Vencida essa etapa inicial, foram desenvolvidas as sub-rotinas para o cálculo das interações  $\pi$  e ligações de hidrogênio, as quais os conceitos serão comentados separadamente.

#### 2.3.5.2 IDENTIFICAÇÃO E LOCALIZAÇÃO DOS ANÉIS AROMÁTICOS

Nas definições das interações  $\pi$  descritas no Item 1.1.1, há a necessidade da localização dos anéis aromáticos existentes na estrutura espacial da molécula. A solução para este problema foi obtida em uma sub-rotina que efetua vetorialmente o cálculo das



distâncias das ligações covalentes entre os átomos da estrutura. Em detalhes, o algoritmo segue as etapas:

- a) Selecionar o primeiro átomo da estrutura;
- b) Efetua-se o mapeamento vetorial para localizar todos os átomos que estão ligados a este covalentemente;
- c) Os pares envolvidos em uma ligação covalente são armazenados em uma tabela temporária;
- d) Repete-se o procedimento “a” até o último átomo da estrutura cristalográfica;
- e) Na tabela temporária, é efetuada uma contagem com o objetivo de identificar os átomos que estão presentes em 4 ou mais ligações covalentes, pois estes são integrantes de um anel;
- f) Utilizando o conceito do Item 1.3.2.1, Figura 1.20; são selecionados 4 átomos que são combinados para formarem 2 vetores, os quais serão utilizados para o cálculo de 2 vetores normais. Caso o ângulo existente entre estes dois vetores normais seja  $0^\circ$ , o anel é plano;
- g) As coordenadas dos átomos integrantes do anel são armazenadas em um banco de dados de anéis planos, conforme Figura 1.21.

Conhecendo-se as coordenadas  $x$ ,  $y$  e  $z$  dos átomos integrantes de um anel, procede-se ao cálculo das coordenadas dos centroides, conforme descrito no Item 1.3.2.1.

### **2.3.5.3 LOCALIZAÇÃO DAS LIGAÇÕES DE HIDROGÊNIO**

Utilizando o mesmo procedimento do Item 2.3.5.3, as ligações covalentes entre o Hidrogênio e demais átomos é utilizada para a identificação das Ligações de hidrogênio, conforme a sequência:

- a) Ordena os átomos da estrutura cristalográfica em ordem alfabética;
- b) Excluir da relação os átomos de Hidrogênio e cria-se uma tabela temporária;

- c) Efetua-se o mapeamento vetorial para localizar todos os átomos que estão ligados a este covalentemente;
- d) Os pares envolvidos em uma ligação covalente são armazenados em uma tabela temporária;
- e) Para cada átomo ligado covalentemente ao Hidrogênio, efetua-se novamente o mapeamento vetorial para localizar os átomos que atendam os parâmetros especificados para a distância  $d_3$  descrita no Item 1.2.1;
- f) Caso existam átomos que atendam à este requisito, é completada a tabela de forma a termos nela o átomo de Hidrogênio, o seu par covalentemente ligado e um terceiro átomo à uma distância  $d_3$  deste;
- g) Por fim, conhecendo-se as coordenadas dos 3 átomos do item “f”, procede-se ao cálculo do ângulo  $\phi$ , conforme o Item 1.3.2.2 e o valor do resultado esteja em conformidade ao especificado na Tabela 1.4, a ligação de hidrogênio é válida;
- h) As coordenadas dos átomos integrantes desta ligação de hidrogênio e as suas propriedades são armazenadas em um banco de dados específico, conforme mostrado na Figura 1.23.

#### **2.3.5.4 FORMATOS NUMÉRICOS**

Possuindo como objetivo usuários de diversos países, o input de dados aceita tanto o padrão numérico brasileiro, quanto o internacional. A diferença entre estes consiste na utilização de vírgulas para a separação das casas decimais e pontos para a separação de milhar no sistema brasileiro. No sistema internacional esses caracteres são invertidos, ficando o ponto como separador decimal e a vírgula para os milhares. Por outro lado, como os visualizadores gráficos utilizam o padrão internacional, os arquivos .mol2 resultantes do processamento seguem o padrão internacional (Caracelli, I et al., 2016)

### 2.3.5.5 TEMPO DE PROCESSAMENTO

Sabendo-se que o tempo de processamento está relacionado à quantidade de átomos existentes em cada molécula analisada, possuímos obrigatoriamente tempos diferentes para cada tipo de molécula. Para avaliação do desempenho do WIM, foram selecionadas quatro estruturas de DNA cujo download foi efetuado diretamente do Protein Data Bank (PDB). Para cada uma das estruturas foi efetuado o mapeamento vetorial para a busca por interações  $\pi$ , utilizando um notebook marca DELL, modelo Inspiron 15R SE, conforme as especificações abaixo:

- Processador: Intel Core i5 3210M;
- Adaptador Gráfico: AMD Radeon HD 7730M;
- Memória RAM: 8 Gb.

Os resultados obtidos foram utilizados na construção da Tabela 2.8.

**Tabela 2.8** Tempo de processamento das estruturas cristalográficas de DNA em um notebook DELL modelo Inspiron 15R SE, extraído de (Caracelli, I. et al; 2016)

PDB ID	Pares de Bases	Quantidade de átomos (sem H)	Possui ligante?	Tempo de Processamento (Segundos)
1BNA	12	566	Não	7,0
2DYW	12	661	No sulco	6,5
1VZK	12	579	No sulco	6,0
1G3X	12	628	Intercalado	7,0

O tempo de processamento obtido para cada uma das estruturas da Tabela 2.8 mostrou ser a principal vantagem do WIM sobre os procedimentos em uso pelos pesquisadores. Empiricamente, o mesmo trabalho quando efetuado manualmente em tela gráfica demandaria 40 horas de trabalho do pesquisador.

### 2.3.5.6 O WIM E A INOVAÇÃO

A proposta original para a criação do WIM surgiu como resultado da dissertação de mestrado intitulada “Mapeamento vetorial de interações eletrostáticas para biomoléculas” (Sacco, 2013) cujo objetivo era preencher uma lacuna existente no mercado, uma vez que na época da sua concepção, o software Discovery Studio Client 3.1<sup>TM</sup> (Accelrys, 2011) e suas versões posteriores efetuam o cálculo das interações  $\pi$ . Todavia, com o uso do mapeamento vetorial foi possível constatar que os cálculos efetuados tanto pelo Discovery Studio Client 3.1<sup>TM</sup> (Accelrys, 2011) como por suas versões posteriores não estão corretos. Tal fato torna o WIM o único software a efetuar de forma automatizada e precisa tanto a localização de interações  $\pi$  como também Ligações de hidrogênio, associando à estas a geração de arquivos para uso estatístico ou outros cálculos matemáticos, além de arquivos para a visualização em *graph sets*.

# Capítulo 3

---

RESULTADOS E DISCUÇÕES

No presente, capítulo utilizando o mapeamento vetorial processado pelo WIM, serão efetuadas análises em uma família de estruturas de DNA com quantidade de pares de bases pré-definida e que apresentam a mesma sequência de nucleotídeos (\*CP\*GP\*CP\*GP\*AP\*AP\*TP\*TP\*CP\*GP\*CP\*G), sendo que cada uma delas será classificada em função da presença ou não de ligantes e caso existam, também por tipo. O mapeamento vetorial permitirá uma melhor compreensão da relação existente entre ligações de hidrogênio e interações  $\pi$  para cada estrutura, a influência dos ligantes na relação entre estas interações e por fim uma nova visão sobre o papel interações  $\pi$  na conformação espacial da dupla hélice do DNA.

### 3 WIM – RESULTADOS OBTIDOS EM ESTRUTURAS DE DNA

Conforme a literatura, quando Watson e Crick (Watson e Crick, 1953) elucidaram a formação da dupla hélice do DNA e os pares formados pelas bases G-C e A-T passaram a ser conhecidos como pares de bases de Watson e Crick (Quigley et al., 1986) e também tidos como os responsáveis pela manutenção da dupla hélice mantendo as duas fitas do DNA em posição via ligações de hidrogênio existentes entre as bases G-C e A-T, conforme a Figura 3.1.

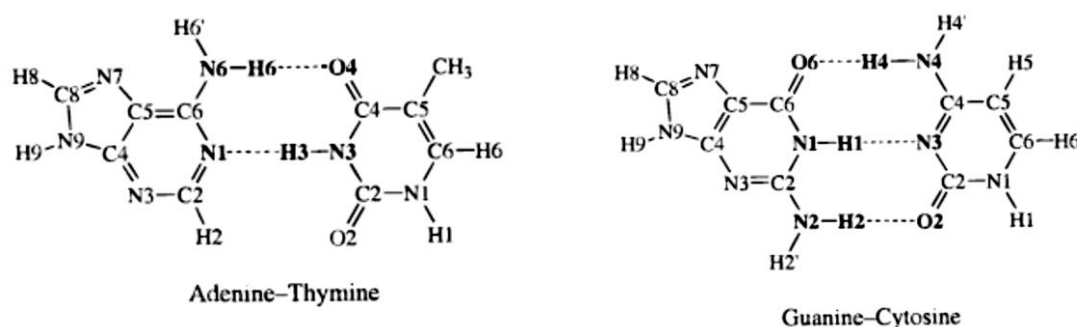


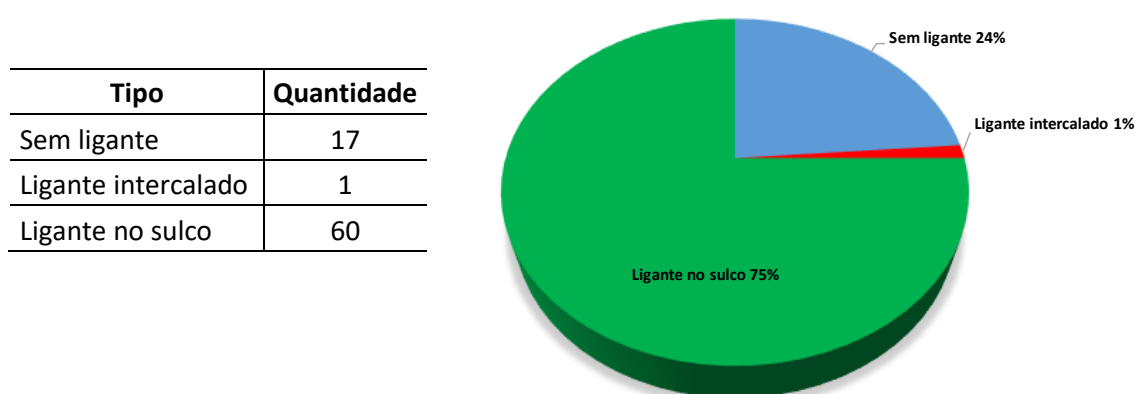
Figura 3.1 Watson Crick, pares de bases, A-T e G-C (Fonseca Guerra et al., 2000).

Complementarmente ao acima exposto, diversos autores (Ma e Dougherty, 1997; Gallivan e Dougherty, 1999; Wintjens et al., 2000; Michael Gromiha et al., 2004; Dougherty, 2007; Du et al., 2012; Pinheiro et al., 2017), argumentam que além das ligações

de hidrogênio referentes aos pares de bases de Watson e Crick, atribuem às interações  $\pi$  existentes entre os anéis paralelos das bases pertencentes à mesma fita como fatores contribuintes para a conformação do DNA. O processamento utilizando o WIM permitiu uma nova visão sobre esse assunto, o qual será melhor explorado a seguir.

### 3.1 DEFINIÇÃO DAS ESTRUTURAS

Considerando-se que o desenvolvimento do mapeamento vetorial (Sacco, 2013) utilizou três estruturas de DNA, descritas no Item 2.1 com sequência de nucleotídeos (\*CP\*GP\*CP\*GP\*AP\*AP\*TP\*TP\*CP\*GP\*CP\*G), optou-se por buscar no PDB (Rscb), PDMSum e NDB (Jersey, 2012) por outras estruturas que apresentassem esta mesma sequência. Na data de 08 de maio de 2016 foram localizados nos três bancos de dados arquivos que possuem a sequência desejada. Em função da presença ou não de ligantes e caso existam, do seu tipo, essas 77 estruturas foram divididas em três grupos e a amostragem em função das suas características é visualizada na Figura 3.2.



**Figura 3.2** Distribuição relativa das estruturas de DNA em função das suas características. Elaborado pelo autor.

A discriminação das estruturas com sequência de nucleotídeos (\*CP\*GP\*CP\*GP\*AP\*AP\*TP\*TP\*CP\*GP\*CP\*G) integrantes da amostragem ilustrada na Figura 4.1 encontra-se relacionada abaixo, conforme mostrado nas Tabelas 3.1 a 3.3.

**Tabela 3.1** Estruturas de DNA sem ligantes, elaborada pelo autor

Cód.	Autor	Ligante	Resolução
SL 1	(Schweitzer et al., 1994)		NMR
SL 2	(Drew et al., 1981)		1.90 Å
SL 3	(Tjandra et al., 2000)		NMR
SL 4	(Chatake et al., 1999)		1.60 Å
SL 5	(Kuszewski et al., 2001)		NMR
SL 6	(Howerton et al., 2001)		1.20 Å
SL 7	(Wu et al., 2003)		NMR
SL 8	(Drew et al., 1982)		2.70 Å
SL 9	(Denisov et al., 1998)		NMR
SL 10	(Shui, Mcfail-Isom, et al., 1998)		1.40 Å
SL 11	(Liu et al., 1998)		3.00 Å
SL 12	(Liu e Subirana, 1999)		1.45 Å
SL 13	(Lercher et al., 2014)		1.32 Å
SL 14	(Holbrook et al., 1985)		1.90 Å
SL 15	(Westhof, 1987)		1.90 Å
SL 16	(Shatzky-Schwartz et al., 1997)		2.50 Å

**Tabela 3.2** Estrutura de DNA com ligante intercalado, elaborada pelo autor

Cód.	Autor	Ligante	Resolução
LI 1	(Malinina et al., 2002)	9ACRIDINE-PEPTIDE DRUG	2.70 Å

Conforme observado acima, para a sequência de nucleotídeos escolhida, há apenas uma com o ligante intercalado (Tabela 3.2).

**Tabela 3.3** Estrutura de DNA com ligantes no sulco, elaborada pelo autor

Cód.	Autor	Ligante	Resolução
LS 1	(Mann et al., 2001)	[3-(4-{2'-[4-(3-DIMETHYLAMINO-PROPOXY)-PHENYL]-3H,3'H-[5,5']BIBENZOIMIDAZOLYL-2-YL}-PHENOXY)-PROPYL]-DIMETHYL-AMINE	2.00 Å
LS 2	(Nunn et al., 1993)	1,3-BIS(AMIDINOPHENOXY)PROPANE	2.10 Å



*Continuação da Tabela 3.3*

LS 3	(Squire, C.J et al., 2000)	2'-(3-iodo-4-methoxyphenyl)-5-(4-methyl-1-piperazinyl)-2,5'-bi-benzimidazole	1.60 Å
LS 4	(Wei et al., 2013)	2-(4'-carbamiimidoylbiphenyl-4-yl)-1H-indole-6-carboximidamide	1.24 Å
LS 5	(Sriram et al., 1992a)	2'-(4-hydroxyphenyl)-5-(4-methyl-1-piperazinyl)-2,5'-bi-benzimidazole	2.00 Å
LS 6	(Simpson, I.J et al., 2000)	2,5-bis{[4-(N-cyclobutylidiaminomethyl)phenyl]}furan	2.00 Å
LS 7	(Neidle e Lee, 2006)	2-{3'-[amino(imino)methyl]biphenyl-4-yl}-1H-benzimidazole-5-carboximidamide	1.65 Å
LS 8	(Larsen et al., 1989)	6-amidine-2-(4-amidino-phenyl)indole	2.40 Å
LS 9	(Adams, A et al., 2005)	1,6-dimethyl-4-(4-(4-(1-methylpyridinium-4-ylamino)phenylcarbamoil)phenylamino)quinolinium	1.80 Å
LS 10	(Sriram et al., 1992a)	2'-(4-ethoxyphenyl)-5-(4-methyl-1-piperazinyl)-2,5'-bi-benzimidazole HOECHST 33342 (Synonym) C27 H28 N6 O	2.25 Å
LS 11	(Quintana, J.R et al., 1991)	2'-(4-hydroxyphenyl)-5-(4-methyl-1-piperazinyl)-2,5'-bi-benzimidazole	2.00 Å
LS 12	(Pjura et al., 1987)	2'-(4-hydroxyphenyl)-5-(4-methyl-1-piperazinyl)-2,5'-bi-benzimidazole	1.90 Å
LS 13	(Quintana, J.R et al., 1991)	2'-(4-hydroxyphenyl)-5-(4-methyl-1-piperazinyl)-2,5'-bi-benzimidazole HOECHST 33258 (Synonym)	1.90 Å
LS 14	(Lee e Neidle, 2006a)	2-(5-{4-[amino(imino)methyl]phenyl}-2-furyl)-1H-benzimidazole-5-carboximidamide	1.63 Å
LS 15	(Squire et al., 1997)	4-[4-[2-amino-4-[4,6-(N-methylquinolinium)amino]benzamido]anilino]-N-methylpyridinium mesylate	2.60 Å
LS 16	(Clark et al., 1997)	5-amidino-2-[2-(4-amidinophenyl)-5-benzimidazolyl]benzimidazole	2.20 Å
LS 17	(Lin et al., 2010)	N',N''-{furan-2,5-diylbis[3-(piperidin-4-yloxy)benzene-4,1-diyl]}dipyridine-2-carboximidamide DB1880	1.90 Å
LS 18	(Edwards et al., 1992)	1,5-bis(4-amidinophenoxy)pentane	2.10 Å
LS 19	(Nunn et al., 1994)	1,5-di(4-amidinophenoxy)-3-oxa-pentane	2.20 Å
LS 20	(Adams, A et al., 2005)	1,6-dimethyl-4-(4-(4-(1-methylpyridinium-4-ylamino)phenylcarbamoil)phenylamino)quinolinium	1.80 Å
LS 21	(Squire, Christopher J et al., 2000)	2'-(3-iodophenyl)-5-(4-methyl-1-piperazinyl)-2,5'-bi-benzimidazole	1.60 Å

Continuação da Tabela 3.3

LS 22	(Squire, Christopher J et al., 2000)	2'-(3-iodophenyl)-5-(4-methyl-1-piperazinyl)-2,5'-bi-benzimidazole	2.10 Å
LS 23	(Martin, R.F et al., 2004)	2'-(3-methyl-4-dimethylaminophenyl)-5-(4-methyl-1-piperazinyl)-2,5'-bi-benzimidazole	2.50 Å
LS 24	(Miao et al., 2005)	2-(4'-amidinobiphenyl-4-yl)-1H-benzimidazole-5-amidine	1.64 Å
LS 25	(Munnur et al., 2012)	2-(4'-carbamimidoyl-2'-hydroxybiphenyl-4-yl)-1H-benzimidazole-5-carboximidamide	1.25 Å
LS 26	(Martin, Roger F et al., 2004)	2'-(4-dimethylaminophenyl)-5-(4-methyl-1-piperazinyl)-2,5'-bi-benzimidazole	2.50 Å
LS 27	(Squire, Christopher J et al., 2000)	2'-(4-dimethylaminophenyl)-5-(4-methyl-1-piperazinyl)-2,5'-bi-benzimidazole	2.20 Å
LS 28	(Quintana, J.R et al., 1991)	2'-(4-hydroxyphenyl)-5-(4-methyl-1-piperazinyl)-2,5'-bi-benzimidazole	2.20 Å
LS 29	(Quintana, Jordi R et al., 1991)	2'-(4-hydroxyphenyl)-5-(4-methyl-1-piperazinyl)-2,5'-bi-benzimidazole	2.20 Å
LS 30	(Teng et al., 1988)	2'-(4-hydroxyphenyl)-5-(4-methyl-1-piperazinyl)-2,5'-bi-benzimidazole	2.25 Å
LS 31	(Mallena et al., 2004)	2-(5-{4-[amino(imino)methyl]phenyl}-2-thienyl)-1H-benzimidazole-6-carboximidamide dihydrochloride	1.77 Å
LS 32	(Zhu et al., 2015)	2,2'-benzene-1,4-diylbis[6-(1,4,5,6-tetrahydropyrimidin-2-yl)-1H-indole]	1.24 Å
LS 33	(Laughton et al., 1996)	2,5-bis(4-guanylphenyl)furan	2.20 Å
LS 34	(Mazur et al., 2000)	2,5-bis-[4-(cyclopenta-1,3-dien-5-ylaminoaminomethyl)-phen-1-yl]furan	2.40 Å
LS 35	(Simpson, I.J et al., 2000)	2,5-bis[4-(N-cyclobutyl-diaminomethyl)phenyl]furan	2.00 Å
LS 36	(Simpson, Ian J et al., 2000)	2,5-bis[4-(N-cyclohexyl-diaminomethyl)phenyl]furan	1.90 Å
LS 37	(Lee e Neidle, 2006b)	2-{4'-[amino(imino)methyl]biphenyl-3-yl}-1H-benzimidazole-6-carboximidamide	2.05 Å
LS 38	(Clark et al., 1996)	3-[5-[5-(4-methyl-piperazin-1-yl)-1H-imidazo[4,5-b]pyridin-2-yl]-benzimidazol-2-yl]-phenol	2.20 Å
LS 39	(Nguyen et al., 2002)	3-[C-[N'-(3-carbamimidoyl-benzylidinium)-hydrazino]-[aminomethylidene]aminium]-iminomethyl]-benzamidinium CGP40215A	1.78 Å
LS 40	(Wei et al., 2013)	4'-(5-carbamimidoyl-1H-benzimidazol-2-yl)biphenyl-4-carboxamide	1.25 Å

Continuação da Tabela 3.3

LS 41	(Neidle et al., 1999)	4-{{[4-HYDROXY-PHENYL]-1H-BENZIMIDAZOLE-5-YL}-BENZIMIDAZOLE-2-YL-[4-HYDROXY-BENZENE] C26 H18 N4 O2	1.80 Å
LS 42	(Wood et al., 1995)	5-(2-IMIDAZOLINYL)-2-[2-(4-HYDROXYPHENYL)-5-BENZIMIDAZOLYL]BENZIMIDAZOLE	2.00 Å
LS 43	(Wei et al., 2013)	6-(1,4,5,6-tetrahydropyrimidin-2-yl)-2-[4'-(1,4,5,6-tetrahydropyrimidin-2-yl)biphenyl-4-yl]-1H-indole	1.27 Å
LS 44	(Campbell et al., 2006)	6-(4,5-DIHYDRO-1H-IMIDAZOL-2-YL)-2-{5-[4-(4,5-DIHYDRO-1H-IMIDAZOL-2-YL)PHENYL]THIEN-2-YL}-1H-BENZIMIDAZOLE	1.36 Å
LS 45	(Adams, Adrienne et al., 2005)	8-METHOXY-1-METHYL-4-(4-(4-(1-METHYLPYRIDINIUM-4-YLAMINO)PHENYL)CARBAMOYL)PHENYLAMINO)QUINOLINIUM	1.60 Å
LS 46	(Brown et al., 1990)	BERENIL DIMINAZINE ACETURATE; 1,3-TRIS-(4'AMIDINOPHENYL)	2.50 Å
LS 47	(Neidle et al., 2006)	BERENIL DIMINAZINE ACETURATE; 1,3-TRIS-(4'AMIDINOPHENYL)	1.65 Å
LS 48	(Sines et al., 2000)	MAGNESIUM ION in minor-groove	1.20 Å
LS 49	(Woods et al., 2000)	MAGNESIUM ION in minor-groove, Mediated by a Conserved Water Cluster	1.82 Å
LS 50	(Wei et al., 2013)	MAGNESIUM ION in minor-groove, Mediated by a Conserved Water Cluster	1.25 Å
LS 51	(Shui, Sines, et al., 1998)	MAGNESIUM ION in minor-groove, Mediated by a Conserved Water Cluster	1.75 Å
LS 52	(Goodsell et al., 1995)	MONOIMIDAZOLE LEXITROPSIN	2.25 Å
LS 53	(Goodsell et al., 1995)	MONOIMIDAZOLE LEXITROPSIN	2.25 Å
LS 54	(Munde et al., 2007)	N,N'-(1H-PYRROLE-2,5-DIYLDI-4,1-PHENYLENE)DIBENZENECARBOXIMIDAMIDE	1.86 Å
LS 55	(Zhu et al., 2015)	N,N'-bis[3-(1,4,5,6-tetrahydropyrimidin-2-yl)phenyl]biphenyl-4,4'-dicarboxamide	1.48 Å
LS 56	(Zhu et al., 2015)	N,N'-bis[3-(4,5-dihydro-1H-imidazol-2-yl)phenyl]biphenyl-4,4'-dicarboxamide	1.31 Å
LS 57	(Sriram et al., 1992b)	NETROPSIN	2.20 Å
LS 58	(Komeda et al., 2006)	PLATINA - TRANS-BIS(HEXYLAMINE)-(DIAMMINE) PLATINUM(II) COMPLEX; (6-AMINOHEXYLAMINE)(TRIAMMINE) PLATINUM(II) COMPLEX spam the minor groove	1.13 Å
LS 59	(Wing et al., 1984)	PLATINUM TRIAMINE ION H9 N3 Pt into the major groove	2.60 Å
LS 60	(Tereshko et al., 1999)	RUBIDIUM ION in minor-groove, Mediated by a Conserved Water Cluster	1.50 Å

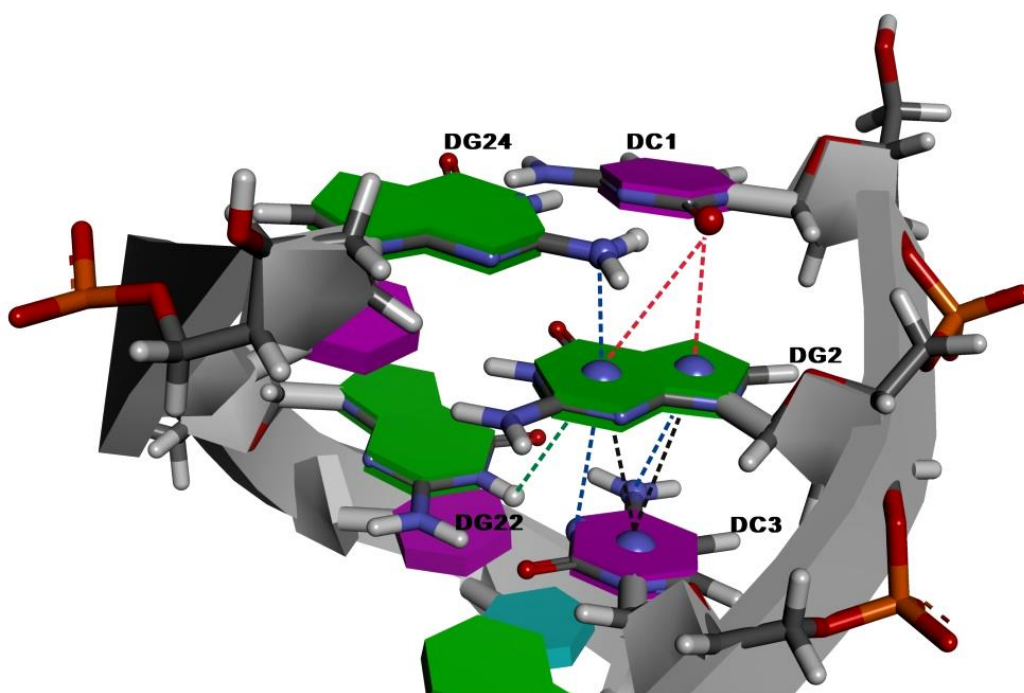
### 3.2 RESULTADOS OBTIDOS

Com o conhecimento obtido tanto no Mapeamento Vetorial para biomoléculas (Sacco, 2013) e também nos trabalhos realizados por Caracelli e colaboradores (Caracelli, I. et al., 2013; Caracelli, I. et al, 2014), onde as operações vetoriais resultaram em novas informações relacionadas às interações não covalentes no DNA, uma vez desenvolvido o WIM (Sacco, A. C. S. et al, 2016), objetivou-se expandir as aplicações já conhecidas para o mapeamento vetorial, utilizando o software em novas análises, tendo em questão as interações  $\pi$  e ligações de hidrogênio. A seguir, os resultados dessas experiências serão apresentados.

### 3.3 PAPEL DAS INTERAÇÕES $\pi$ NA CONFORMAÇÃO DA ESTRUTURA DO DNA

A literatura utilizada nessa pesquisa apresenta dois pontos em comum, o primeiro deles é que a estrutura da dupla hélice característica do DNA é mantida pelas ligações de hidrogênio já anteriormente citadas (Watson e Crick, 1953; Quigley et al., 1986; Kielkopf et al., 1998; Williams e Maher III, 2000; Bao, 2002), o outro ponto comum é que interações  $\pi$  também atuam na conformação estrutural do DNA conforme Guerra e colaboradores (Fonseca Guerra et al., 2000) e demais outros autores (Mignon et al., 2005; Ermondi e Caron, 2006; Tewari e Dubey, 2008; Seng et al., 2010; Takahashi et al., 2010). Tewari e Dubey (2008) definiram que as interações  $\pi \cdots \pi$  e  $C-H \cdots \pi$  são as principais responsáveis pelo empilhamento das bases. Em suma, segundo a pesquisa efetuada, conclui-se que cada fita tem o formato de hélice decorrente das interações  $\pi$  e também pelas ligações de hidrogênio, sendo que estas últimas mantêm as duas fitas unidas, formando a dupla hélice. As Figuras 2.7, 2.8 e 2.9 do Capítulo 2 corroboram o papel das interações  $\pi$  atuando entre as bases de uma mesma fita, porém, a Figura 3.3 revela a existência de interações entre bases de fitas diferentes, possuindo estas interações origem no centroide do anel de 6 átomos da Guanina 2 na Fita A e tendo como destinos um

nitrogênio na Guanina 24 e um hidrogênio na Guanina 22, ambos os destinos na Fita B do DNA analisado.



**Figura 3.3** Representação gráfica de parte de uma molécula de DNA (Drew *et al.*, 1981) com as interações  $\pi$  existentes entre as bases, determinadas vetorialmente. Utilizado o software Discovery Studio 3.1™, extraído de Sacco (2013)

A informação contida na imagem acima, deu origem à uma primeira questão: existem interações  $\pi$  entre as bases de fitas diferentes de forma a contribuir para a manutenção da dupla hélice e, caso positivo, poderia haver um padrão específico para elas? Dessa forma, foi utilizado o WIM para localizar as interações  $\pi$  e as ligações de hidrogênio para cada uma das 77 estruturas relacionadas nas Tabelas 3.1, 3.2 e 3.3, a Tabela 3.4 mostra a estrutura de dados para interações  $\pi$  obtidas como output desse processamento, sendo esse arquivo compatível para a utilização na planilha eletrônica Microsoft Excel (Microsoft.Com, 2012).

**Tabela 3.4** Interações  $\pi$  entre fitas mapeadas pelo WIM para estrutura de DNA sem ligante (Drew *et al.*, 1981), sendo *Distance* expressa em (Å) e o ângulo  $\alpha$  em ( $^{\circ}$ ), tabela elaborada pelo autor.

Origin		Destination		Distance	$\alpha$
DG2	Ct6	DG22	H1	3,7788	27,712
DG2	Ct6	DG24	N2	3,6022	18,3243
DG2	Ct6	DG24	H22	3,4728	8,0744
DG4	Ct6	DG22	H22	3,4977	20,313
DG10	Ct6	DG14	O6	3,6305	28,6492
DG10	Ct6	DG16	H1	3,7167	28,6692
DG10	Ct6	DG16	H22	3,4156	28,2835
DG12	Ct6	DG14	H22	3,693	23,8587
DG16	Ct6	DG10	H22	3,574	24,5559
DG22	Ct6	DG4	H1	3,5274	28,7619
DG22	Ct6	DG4	H22	3,5428	27,9633
DG24	Ct6	DG2	N1	4,0697	29,4

Como pode ser constatado na tabela acima, o WIM além de localizar as interações  $\pi$ , fornece informações sobre cada uma delas, tais como a sua direcionalidade, quantidade de átomos do anel aromático de origem, átomo ou centroide de destino e sua localização na estrutura e os parâmetros geométricos definidos pela literatura (Tabela 1.2). A disposição dos dados obtidos para todas as estruturas no formato de tabela, permite que sejam analisados utilizando ferramentas de planilha eletrônica, como por exemplo, o Microsoft Excel (Microsoft.Com, 2012). A seguir, serão mostrados os resultados dessa análise para cada um dos grupos definidos nas Tabelas 3.1, 3.2 e 3.3.

### 3.3.1 ESTRUTURAS DE DNA SEM LIGANTES

O WIM, ao processar as estruturas integrantes da Tabela 3.1, identificou um total de 268 interações entre bases de fitas diferentes. Ao serem analisadas, são obtidas informações que servem como base para a avaliação de um padrão comum das interações  $\pi$  para estas estruturas. As Tabelas de 3.5 a 3.12 possibilitam uma visão qualitativa e quantitativas das interações entre fitas de estruturas sem ligantes.

**Tabela 3.5** Interações  $\pi$  entre fitas mapeadas pelo WIM para as estruturas de DNA sem ligantes, quantificação e localização das interações entre fitas, possuindo como destino átomos da Guanina 2, tabela elaborada pelo autor.

Destino >	DG2				
Origem (localização do centroide: CT6)	H1	H21	H22	N1	N2
DG2					
DG4					
DG10					
DG12					
DG14					
DG16					
DG22	5				
DG24	5	3	4	7	5
Total Geral	10	3	4	7	5

**Tabela 3.6** Interações  $\pi$  entre fitas mapeadas pelo WIM para as estruturas de DNA sem ligantes, quantificação e localização das interações entre fitas, possuindo como destino átomos e centroide da Guanina 4, tabela elaborada pelo autor.

Destino >	DG4						
Origem (localização do centroide: CT6)	H1	H21	H22	N1	N2	Ct6	N3
DG2							
DG4							
DG10							
DG12							
DG14							
DG16							
DG22	6	5	4	3	6	2	1
DG24							
Total Geral	6	5	4	3	6	2	1

**Tabela 3.7** Interações  $\pi$  entre fitas mapeadas pelo WIM para as estruturas de DNA sem ligantes, quantificação e localização das interações entre fitas, possuindo como destino átomos e centroide da Guanina 10, tabela elaborada pelo autor.

Destino >	DG10					
Origem (localização do centroide: CT6)	H1	H21	H22	N1	N2	Ct6
DG2						
DG4						
DG10						
DG12						
DG14	3					
DG16	6	3	5	5	2	1
DG22						
DG24						
Total Geral	9	3	5	5	2	1

**Tabela 3.8** Interações  $\pi$  entre fitas mapeadas pelo WIM para as estruturas de DNA sem ligantes, quantificação e localização das interações entre fitas, possuindo como destino átomos e centroide da Guanina 12, tabela elaborada pelo autor.

Destino >	DG12					
Origem (localização do centroide: CT6)	H1	H21	H22	N1	N2	Ct6
DG2						
DG4						
DG10						
DG12						
DG14	6	4	5	4	5	1
DG16						
DG22						
DG24						
Total Geral	6	4	5	4	5	1

**Tabela 3.9** Interações  $\pi$  entre fitas mapeadas pelo WIM para as estruturas de DNA sem ligantes, quantificação e localização das interações entre fitas, possuindo como destino átomos da Guanina 14, tabela elaborada pelo autor.

Destino >	DG14					
Origem (localização do centroide: CT6)	H1	H21	H22	N1	N2	O6
DG2						
DG4						
DG10						5
DG12	3	4	4	3	4	
DG14						
DG16						
DG22						
DG24						
Total Geral	3	4	4	3	4	5

**Tabela 3.10** Interações  $\pi$  entre fitas mapeadas pelo WIM para as estruturas de DNA sem ligantes, quantificação e localização das interações entre fitas, possuindo como destino átomos e centroide da Guanina 16, tabela elaborada pelo autor.

Destino >	DG16						
Origem (localização do centroide: CT6)	H1	H21	H22	N1	N2	Ct6	N3
DG2							
DG4							
DG10	5	5	4	3	6	2	1
DG12							
DG14							
DG16							
DG22							
DG24							
Total Geral	5	5	4	3	6	2	1



**Tabela 3.11** Interações  $\pi$  entre fitas mapeadas pelo WIM para as estruturas de DNA sem ligantes, quantificação e localização das interações entre fitas, possuindo como destino átomos e centroide da Guanina 22, tabela elaborada pelo autor.

Destino >	DG22					
	H1	H21	H22	N1	N2	Ct6
Origem (localização do centroide: CT6)						
DG2	6					
DG4	5	3	5	3	3	1
DG10						
DG12						
DG14						
DG16						
DG22						
DG24						
Total Geral	11	3	5	3	3	1

**Tabela 3.12** Interações  $\pi$  entre fitas mapeadas pelo WIM para as estruturas de DNA sem ligantes, quantificação e localização das interações entre fitas, possuindo como destino átomos e centroide da Guanina 24, tabela elaborada pelo autor.

Destino >	DG24					
	H1	H21	H22	N1	N2	Ct6
Origem (localização do centroide: CT6)						
DG2	6	4	6	7	7	2
DG4						
DG10						
DG12						
DG14						
DG16						
DG22						
DG24						
Total Geral	6	4	6	7	7	2

Como pode ser visto nas Tabelas 3.5 a 3.12, para o grupo de estruturas sem ligantes, as interações  $\pi$  ocorrem especificamente entre centroides e átomos das guaninas DG2, DG4, DG10, DG12, DG14, DG16, DG22 e DG24. Para uma melhor compreensão, as Tabelas 3.13 e 3.14 consolidam as interações mapeadas por guaninas.

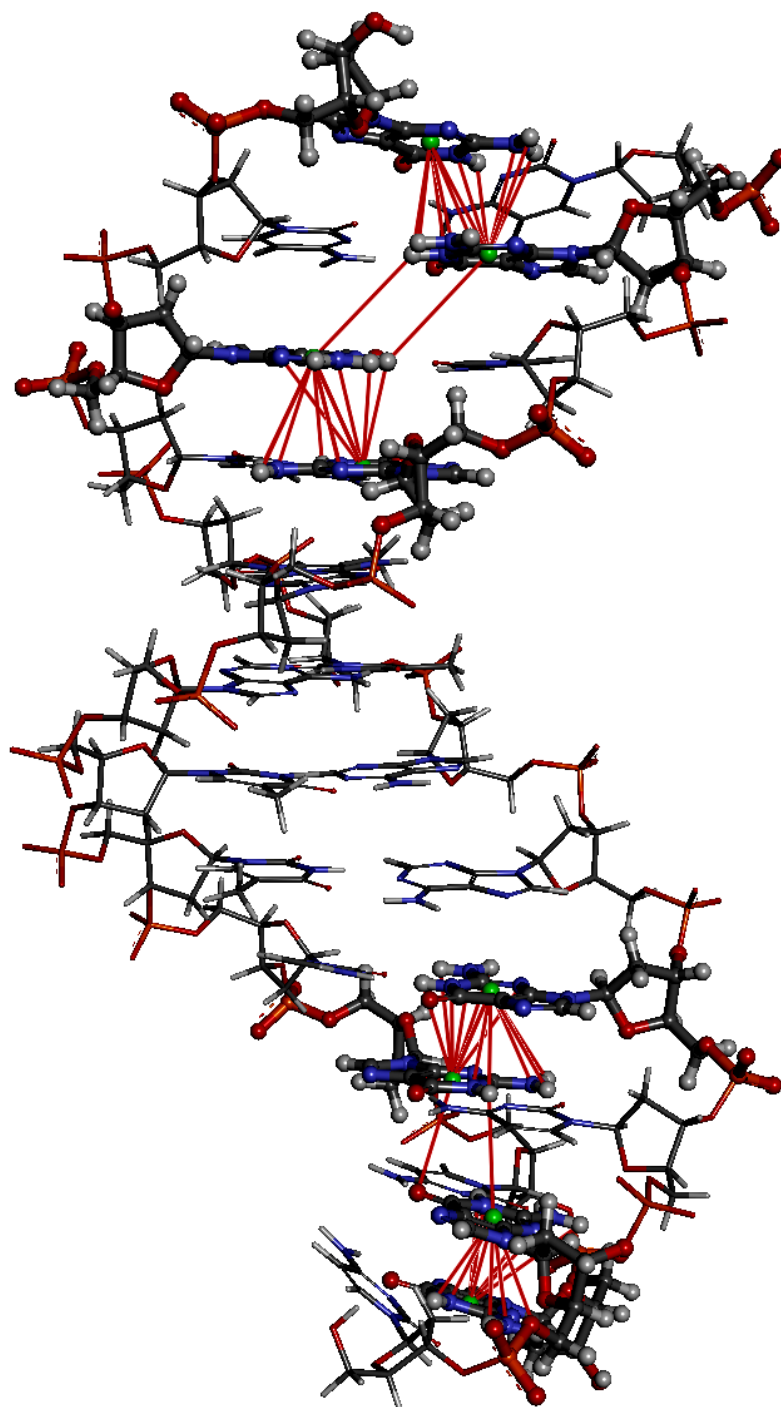
**Tabela 3.13** Distribuição em valores absolutos das interações  $\pi$  entre fitas mapeadas pelo WIM para as estruturas de DNA sem ligantes, evidenciando as ligações existentes entre fitas diferentes, tabela elaborada pelo autor.

Origem	Destino								Total por origem
	DG2	DG4	DG10	DG12	DG14	DG16	DG22	DG24	
DG2							10	41	51
DG4							23		23
DG10					7	28			35
DG12					22				22
DG14			3	31					34
DG16			27						27
DG22	9	33							42
DG24	34								34
Total por destino	43	33	30	31	29	28	33	41	268

**Tabela 3.14** Distribuição em valores relativos das interações  $\pi$  entre fitas mapeadas pelo WIM para as estruturas de DNA sem ligantes, evidenciando as ligações existentes entre fitas diferentes, tabela elaborada pelo autor.

Origem	Destino								Total por origem
	DG2	DG4	DG10	DG12	DG14	DG16	DG22	DG24	
DG2							4%	15%	19%
DG4							9%		9%
DG10					3%	10%			13%
DG12					8%				8%
DG14			1%	12%					13%
DG16			10%						10%
DG22	3%	12%							16%
DG24	13%								13%
Total por destino	16%	12%	11%	12%	11%	10%	12%	15%	100%

Observando a Tabela 3.14, constata-se que há uma distribuição equânime destas interações entre guaninas específicas. Como o WIM também gerou arquivos para uso em visualizadores gráficos, as interações acima reportadas foram transformadas em imagem e mostradas na Figura 3.4.



**Figura 3.4** Interações  $\pi$  existentes entre as bases de fitas diferentes mapeadas em estruturas sem ligantes, elaborado pelo autor.

Pode-se constatar pela imagem acima, que há uma concentração de interações  $\pi$  entre fitas nos extremos da estrutura.

### 3.3.2 ESTRUTURA DE DNA COM LIGANTE INTERCALADO

Para que fosse possível repetir a mesma análise para o único DNA com a sequência específica de nucleotídeos (\*CP\*GP\*CP\*GP\*AP\*AP\*TP\*TP\*CP\*GP\*CP\*G) que possui ligante intercalado (Malinina *et al.*, 2002), foram seguidos os mesmos procedimentos utilizados anteriormente com o grupo de estruturas sem ligantes, conforme visto nas Tabelas de 3.15 a 3.22.

**Tabela 3.15** Interações  $\pi$  entre fitas mapeadas pelo WIM para as estruturas de DNA com ligante intercalado, quantificação e localização das interações entre fitas, possuindo como destino átomos da Guanina 2, tabela elaborada pelo autor.

Destino >	DG2	
Origem (localização do centroide: CT6)	H1	N1
DG2		
DG4		
DG10		
DG12		
DG24	1	1
Total Geral	1	1

**Tabela 3.16** Interações  $\pi$  entre fitas mapeadas pelo WIM para as estruturas de DNA com ligante intercalado, quantificação e localização das interações entre fitas, possuindo como destino átomos da Guanina 14, tabela elaborada pelo autor.

Destino >	DG14	
Origem (localização do centroide: CT6)	H1	H21
DG2		
DG4		
DG10	1	
DG12		1
DG24		
Total Geral		

**Tabela 3.17** Interações  $\pi$  entre fitas mapeadas pelo WIM para as estruturas de DNA com ligante intercalado, quantificação e localização da interação entre fitas, possuindo como destino átomo da Guanina 16, tabela elaborada pelo autor.

Destino >	DG16
Origem (localização do centroide: CT6)	H21
DG2	
DG4	
DG10	1
DG12	
DG24	
Total Geral	1

**Tabela 3.18** Interações  $\pi$  entre fitas mapeadas pelo WIM para as estruturas de DNA com ligante intercalado, quantificação e localização da interação entre fitas, possuindo como destino átomo da Guanina 22, tabela elaborada pelo autor.

Destino >	DG22
Origem (localização do centroide: CT6)	H22
DG2	
DG4	1
DG10	
DG12	
DG24	
Total Geral	1

**Tabela 3.19** Interações  $\pi$  entre fitas mapeadas pelo WIM para as estruturas de DNA com ligante intercalado, quantificação e localização das interações entre fitas, possuindo como destino átomos da Guanina 24, tabela elaborada pelo autor.

Destino >	DG24		
Origem (localização do centroide: CT6)	H21	H22	N3
DG2	1	1	1
DG4			
DG10			
DG12			
DG24			
Total Geral	1	1	1

Como pode ser visto nas Tabelas 3.15 a 3.19, a notável redução das interações entre fitas já foi anteriormente abordada por outros pesquisadores (Berman e Young, 1981; Shaikh e Jayaram, 2012), os quais haviam concluído que o ligante intercalado promove alterações físicas no DNA. Observa-se que, embora as mesmas guaninas DG2, DG4, DG10, DG12, DG14, DG16, DG22 e DG24 estejam envolvidas em interações  $\pi$  entre fitas, há uma visível alteração na relação origem-destino das interações, conforme mostrado nas Tabelas 3.20 e 3.21.

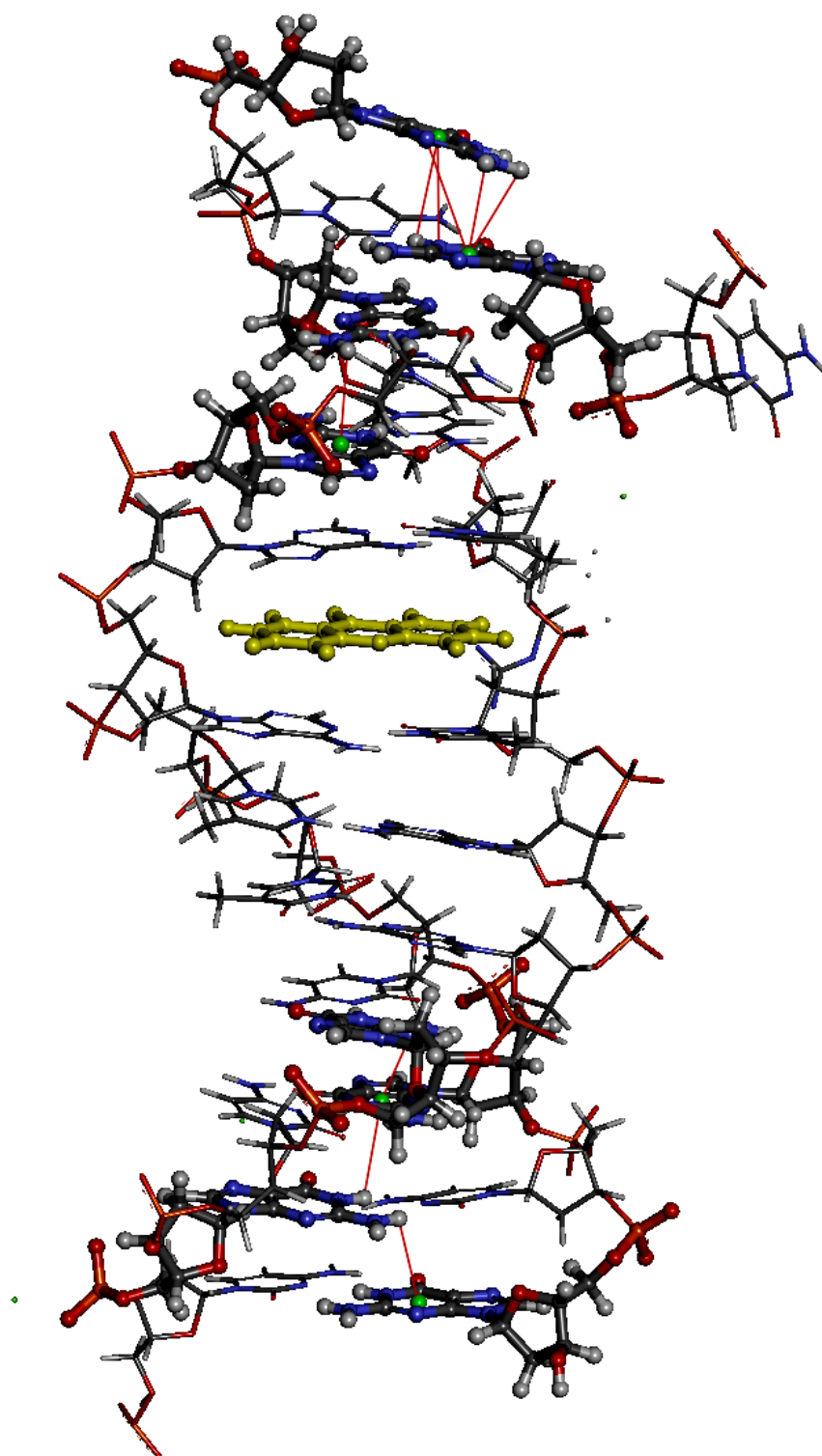
**Tabela 3.20** Distribuição em valores absolutos das interações  $\pi$  entre fitas mapeadas pelo WIM para a estrutura de DNA com ligante intercalado, evidenciando as ligações existentes entre fitas diferentes, tabela elaborada pelo autor.

Origem	Destino					Total por origem
	DG2	DG14	DG16	DG22	DG24	
DG2					3	3
DG4				1		1
DG10		1	1			2
DG12		1				1
DG24	2					2
Total por destino	2	2	1	1	3	9

**Tabela 3.21** Distribuição em valores relativos das interações  $\pi$  entre fitas mapeadas pelo WIM para a estrutura de DNA com ligante intercalado, evidenciando as ligações existentes entre fitas diferentes, tabela elaborada pelo autor.

Origem	Destino					Total por origem
	DG2	DG14	DG16	DG22	DG24	
DG2					100%	33%
DG4				100%		11%
DG10		50%	100%			22%
DG12		50%				11%
DG24	100%					22%
Total por destino	100%	100%	100%	100%	100%	100%

Utilizando os arquivos gerados pelo WIM, foram plotadas as interações entre fitas do DNA mapeadas (Figura 3.5).



**Figura 3.5** Interações  $\pi$  existentes entre as bases de fitas diferentes mapeadas em estrutura com ligante intercalado (Malinina et al., 2002), elaborado pelo autor.

### 3.3.3 ESTRUTURAS DE DNA COM LIGANTES NO SULCO

Complementando a análise das interações  $\pi$  entre as fitas do DNA, abaixo, as Tabelas 3.22 a 3.29 mostram os resultados obtidos.

**Tabela 3.22** Interações  $\pi$  entre fitas mapeadas pelo WIM para as estruturas de DNA com ligantes no sulco, quantificação e localização das interações entre fitas, possuindo como destino átomos da Guanina 2, tabela elaborada pelo autor.

Destino >	DG2					
Origem (localização do centroide: CT6)	H1	H21	H22	N1	N2	O6
DG2						
DG4						
DG10						
DG12						
DG14						
DG16						
DG22	39					1
DG24	35	3	42	27	15	1
Total Geral	74	3	42	27	15	2

**Tabela 3.23** Interações  $\pi$  entre fitas mapeadas pelo WIM para as estruturas de DNA com ligantes no sulco, quantificação e localização das interações entre fitas, possuindo como destino átomos e centroide da Guanina 4, tabela elaborada pelo autor.

Destino >	DG4						
Origem (localização do centroide: CT6)	H1	H21	H22	N1	N2	Ct6	N3
DG2							
DG4							
DG10							
DG12							
DG14							
DG16							
DG22	24	5	48	3	6	2	1
DG24							
Total Geral	24	5	48	3	6	2	1



**Tabela 3.24** Interações  $\pi$  entre fitas mapeadas pelo WIM para as estruturas de DNA com ligantes no sulco, quantificação e localização das interações entre fitas, possuindo como destino átomos e centroide da Guanina 10, tabela elaborada pelo autor.

Destino >	DG10					
Origem (localização do centroide: CT6)	H1	H21	H22	N1	N2	Ct6
DG2						
DG4						
DG10						
DG12						
DG14	36		1			
DG16	6	3	33	5	2	1
DG22						
DG24						
Total Geral	42	3	34	5	2	1

**Tabela 3.25** Interações  $\pi$  entre fitas mapeadas pelo WIM para as estruturas de DNA com ligantes no sulco, quantificação e localização das interações entre fitas, possuindo como destino átomos e centroide da Guanina 12, tabela elaborada pelo autor.

Destino >	DG12					
Origem (localização do centroide: CT6)	H1	H21	H22	N1	N2	Ct6
DG2						
DG4						
DG10						
DG12						
DG14	11	5	48	12	31	1
DG16						
DG22						
DG24						
Total Geral	11	5	48	12	31	1

**Tabela 3.26** Interações  $\pi$  entre fitas mapeadas pelo WIM para as estruturas de DNA com ligantes no sulco, quantificação e localização das interações entre fitas, possuindo como destino átomos da Guanina 14, tabela elaborada pelo autor.

Destino >	DG14					
Origem (localização do centroide: CT6)	H1	H21	H22	N1	N2	O6
DG2						
DG4						
DG10	11			3		38
DG12	4	4	31	6	6	
DG14						
DG16						
DG22						
DG24						
Total Geral	15	4	31	9	6	38

**Tabela 3.27** Interações  $\pi$  entre fitas mapeadas pelo WIM para as estruturas de DNA com ligantes no sulco, quantificação e localização das interações entre fitas, possuindo como destino átomos e centroide da Guanina 16, tabela elaborada pelo autor.

Destino >	DG16					
Origem (localização do centroide: CT6)	H1	H21	H22	N1	N2	Ct6
DG2						
DG4						
DG10	6	5	18	3	6	2
DG12						
DG14						
DG16						
DG22						
DG24						
Total Geral	6	5	18	3	6	2

**Tabela 3.28** Interações  $\pi$  entre fitas mapeadas pelo WIM para as estruturas de DNA com ligantes no sulco, quantificação e localização das interações entre fitas, possuindo como destino átomos e centroide da Guanina 22, tabela elaborada pelo autor.

Destino >	DG22					
Origem (localização do centroide: CT6)	H1	H21	H22	N1	N2	Ct6
DG2	42					
DG4	5	3	48	3	5	1
DG10						
DG12						
DG14						
DG16						
DG22						
DG24						
Total Geral	47	3	48	3	5	1

**Tabela 3.29** Interações  $\pi$  entre fitas mapeadas pelo WIM para as estruturas de DNA com ligantes no sulco, quantificação e localização das interações entre fitas, possuindo como destino átomos e centroide da Guanina 24, tabela elaborada pelo autor.

Destino >	DG24					
Origem (localização do centroide: CT6)	H1	H21	H22	N1	N2	Ct6
DG2	16	4	53	18	51	3
DG4						
DG10						
DG12						
DG14						
DG16						
DG22						
DG24						
Total Geral	16	4	53	18	51	3

As Tabelas 3.22 a 3.19, mostram que as 843 interações  $\pi$  mapeadas nas estruturas com ligante no sulco apresentam um comportamento semelhante às das estruturas sem ligante, onde estas ocorrem entre centroides e átomos das guaninas DG2, DG4, DG10, DG12, DG14, DG16, DG22 e DG24. As Tabelas 3.30 e 3.31 mostram essa semelhança.

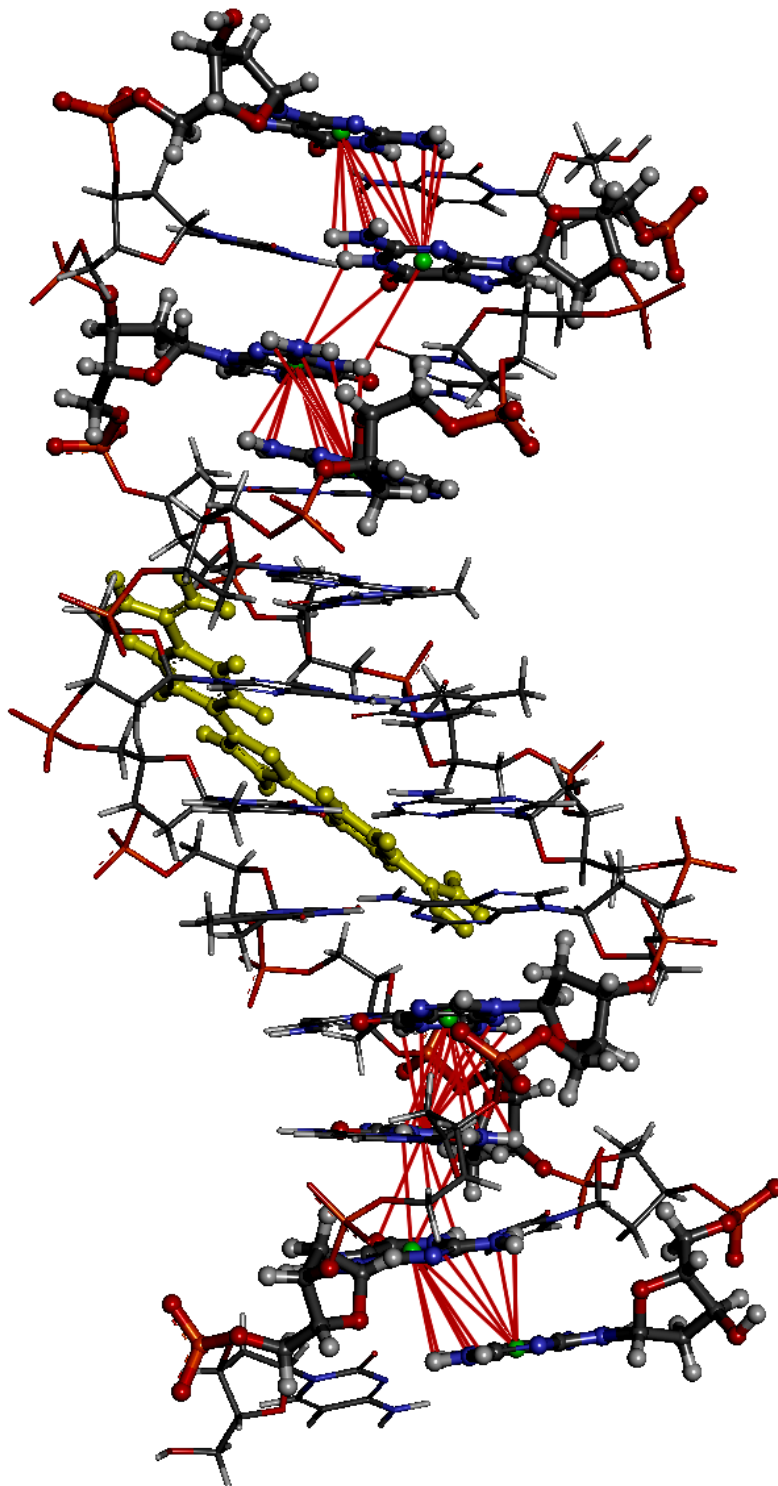
**Tabela 3.30** Distribuição em valores absolutos das interações  $\pi$  entre fitas mapeadas pelo WIM para as estruturas de DNA com ligantes no sulco, evidenciando as ligações existentes entre fitas diferentes, tabela elaborada pelo autor.

Origem	Destino								Total por origem
	DG2	DG4	DG10	DG12	DG14	DG16	DG22	DG24	
DG2							42	145	187
DG4							65		65
DG10					52	41			93
DG12					51				51
DG14			37	108					145
DG16			50						50
DG22	40	89							129
DG24	123								123
Total por destino	163	89	87	108	103	41	107	145	843

**Tabela 3.31** Distribuição em valores relativos das interações  $\pi$  entre fitas mapeadas pelo WIM para as estruturas de DNA com ligantes no sulco, evidenciando as ligações existentes entre fitas diferentes, tabela elaborada pelo autor.

Origem	Destino								Total por origem
	DG2	DG4	DG10	DG12	DG14	DG16	DG22	DG24	
DG2							5%	17%	22%
DG4							8%		8%
DG10					6%	5%			11%
DG12					6%				6%
DG14			4%	13%					17%
DG16			6%						6%
DG22	5%	11%							15%
DG24	15%								15%
Total por destino	19%	11%	10%	13%	12%	5%	13%	17%	100%

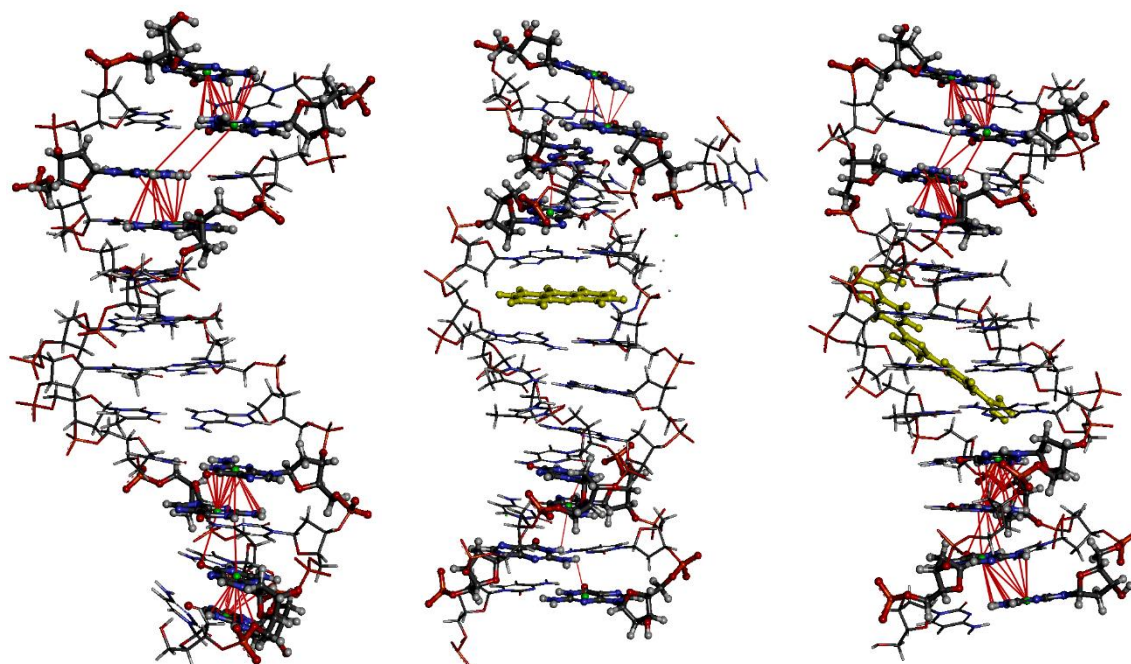
Da mesma forma que as Figuras 3.4 e 3.5, a Figura 3.6 exemplifica uma estrutura padrão com ligantes.



**Figura 3.6** Interações  $\pi$  existentes entre as bases de fitas diferentes mapeadas em estruturas com ligantes no sulco, elaborado pelo autor.

### 3.3.4 VISÃO GERAL DAS INTERAÇÕES $\pi$

Consolidando os dados obtidos e mostrados nos Itens 3.3.1, 3.3.2 e 3.3.3, podemos concluir que para as estruturas de DNA com a sequência de nucleotídeos (\*CP\*GP\*CP\*GP\*AP\*AP\*TP\*TP\*CP\*GP\*CP\*G), existe um padrão na ocorrência de interações  $\pi$  entre fitas, sendo que estas sempre possuem a sua origem no centroide de do anel de 6 átomos de uma guanina (Ct6) e tem como destinos específicos os átomos mapeadas H1, H21, H22, N1, N2 e O6, além do Ct6 das guaninas DG2, DG4, DG10, DG12, DG14, DG16, DG22 2 DG24. A Figura 3.7 mostra um comparativo dessas interações entre as estruturas analisadas e a relação entre as bases de diferentes fitas envolvidas em interações  $\pi$  é mostrada nas Tabelas 3.32 e 3.33.



**Figura 3.7** Imagens das estruturas de DNA com as interações  $\pi$  entre fitas identificadas por mapeamento vetorial. Da esquerda para a direita: a) estrutura sem ligante, b) estrutura com ligante intercalado e c) estrutura com ligante no sulco. Elaborado pelo autor.

**Tabela 3.32** Distribuição em valores absolutos das interações  $\pi$  entre fitas mapeadas pelo WIM para as estruturas de DNA com sequência de nucleotídeos (\*CP\*GP\*CP\*GP\*AP\*AP\*TP\*TP\*CP\*GP\*CP\*G), evidenciando as ligações existentes entre fitas diferentes, tabela elaborada pelo autor.

Origem	Destino								Total por origem
	DG2	DG4	DG10	DG12	DG14	DG16	DG22	DG24	
DG2							52	174	226
DG4							75		75
DG10					57	48			105
DG12					64				64
DG14			37	129					166
DG16			61						61
DG22	50	104							154
DG24	153								153
Total por destino	203	104	98	129	121	48	127	174	1004

**Tabela 3.33** Distribuição em valores relativos das interações  $\pi$  entre fitas mapeadas pelo WIM para as estruturas de DNA com sequência de nucleotídeos (\*CP\*GP\*CP\*GP\*AP\*AP\*TP\*TP\*CP\*GP\*CP\*G), evidenciando as ligações existentes entre fitas diferentes, tabela elaborada pelo autor.

Origem	Destino								Total por origem
	DG2	DG4	DG10	DG12	DG14	DG16	DG22	DG24	
DG2							5%	17%	23%
DG4							7%		7%
DG10					6%	5%			10%
DG12					6%				6%
DG14			4%	13%					17%
DG16			6%						6%
DG22	5%	10%							15%
DG24	15%								15%
Total por destino	20%	10%	10%	13%	12%	5%	13%	17%	100%

### 3.4 INTERAÇÕES ENTRE OS LIGANTES E AS ESTRUTURAS DO DNA

Continuando com a análise dos dados fornecidos pelo WIM, surgiu um segundo questionamento: quais os tipos de interações que ocorrem entre o DNA e os seus ligantes? Várias teorias abordam os mecanismos pelos quais os ligantes incorporam-se ao DNA, todavia, como o WIM efetua o seu processamento com base em padrões geométricos e utilizando cálculo vetorial em estruturas cristalográficas, os resultados por ele obtidos evidenciam o comportamento do ligante em relação à estrutura do DNA e vice-versa. Para efeito de avaliação, foram consideradas que estas ligações entre os complexos e DNA sejam intermediadas apenas pelas interações  $\pi$  e ligações de hidrogênio. Dessa forma, com os arquivos resultantes do processamento do grupo de estruturas de DNA com ligante intercalado e ligantes no sulco, efetuou-se um estudo comparativo entre as interações  $\pi$  e ligações de hidrogênio envolvidas na relação DNA e ligantes e ainda onde está a origem desta ligação para cada estrutura. Com base nesse pressuposto, algumas estruturas relacionadas na Tabela 3.3 foram tratadas a parte, pois os ligantes envolvidos não possuem anéis aromáticos. A Tabela 3.34 mostra a relação existente entre as interações  $\pi$  e ligações de hidrogênio e as origens destas para a estrutura LI 1 (Malinina et al., 2002).

**Tabela 3.34** Relação existente entre as interações  $\pi$  e ligações de hidrogênio e as origens destas para a estrutura com o ligante intercalado, tabela elaborada pelo autor.

Cód.	Interações $\pi$				Ligação de Hidrogênio				Relação Interação $\pi$ x Ligação de Hidrogênio
	Totais	Origem no DNA	Origem no ligante	% no ligante	Totais	Origem no DNA	Origem no ligante	% no ligante	
LI 1	31	15	16	51,61%	0	0	0	0,00%	-
Total geral	31	15	16	51,61%	0	0	0	0,00%	-

Vale ressaltar que o caso do ligante intercalado para o grupo de estruturas de DNA com a sequência de nucleotídeos (\*CP\*GP\*CP\*GP\*AP\*AP\*TP\*TP\*CP\*GP\*CP\*G), é único e também a direcionalidade existente nas interações  $\pi$  entre ligante e DNA, devendo assim ser tratado como uma exceção. Conforme dito anteriormente, algumas estruturas relacionadas com ligante no sulco foram tratadas a parte, pois os seus ligante são íons mantidos no sulco menor pela ação de uma cadeia de moléculas de água (Shui, Sines, et

al., 1998; Tereshko et al., 1999; Sines et al., 2000; Woods et al., 2000; Wei et al., 2013). Estas estruturas de DNA com íons sequestrados por moléculas de água no sulco menor estão mostradas na Tabela 3.35.

**Tabela 3.35** Estruturas de DNA com íons sequestrados por moléculas de água no sulco menor, tabela elaborada pelo autor.

Cód.	Interações $\pi$				Ligação de Hidrogênio				Relação Interação $\pi$ x Ligação de Hidrogênio
	Totais	Origem no DNA	Origem no ligante	% no ligante	Totais	Origem no DNA	Origem no ligante	% no ligante	
LS 48	0	0	0	0,00%	0	0	0	0,00%	0,00%
LS 49	0	0	0	0,00%	0	0	0	0,00%	0,00%
LS 50	0	0	0	0,00%	0	0	0	0,00%	0,00%
LS 51	0	0	0	0,00%	0	0	0	0,00%	0,00%
LS 60	0	0	0	0,00%	0	0	0	0,00%	0,00%
Total geral	0	0	0	0,00%	0	0	0	0,00%	0,00%

Ainda dentro do rol das estruturas tratadas como exceção, existem outras duas, cujos compostos à base de platina envolvidos não possuem anéis aromáticos (Wing et al., 1984; Komeda et al., 2006). A Tabela 3.36 exhibe as ligações de hidrogênio mapeadas.

**Tabela 3.36** Estruturas de DNA com compostos à base de platina, tabela elaborada pelo autor.

Cód.	Interações $\pi$				Ligação de Hidrogênio				Relação Interação $\pi$ x Ligação de Hidrogênio
	Totais	Origem no DNA	Origem no ligante	% no ligante	Totais	Origem no DNA	Origem no ligante	% no ligante	
LS 58	0	0	0	0,00%	2	0	2	100,00%	0,00%
LS 59	0	0	0	0,00%	1	0	1	100,00%	0,00%
Total geral	0	0	0	0,00%	3	0	3	100,00%	0,00%

Uma vez especificadas nas Tabelas 3.35 e 3.36 as exceções encontradas pelo WIM nas estruturas selecionadas, foram relacionadas para análise as 53 estruturas cujos ligantes no sulco possuem anéis aromáticos, conforme mostra a Tabela 3.37.

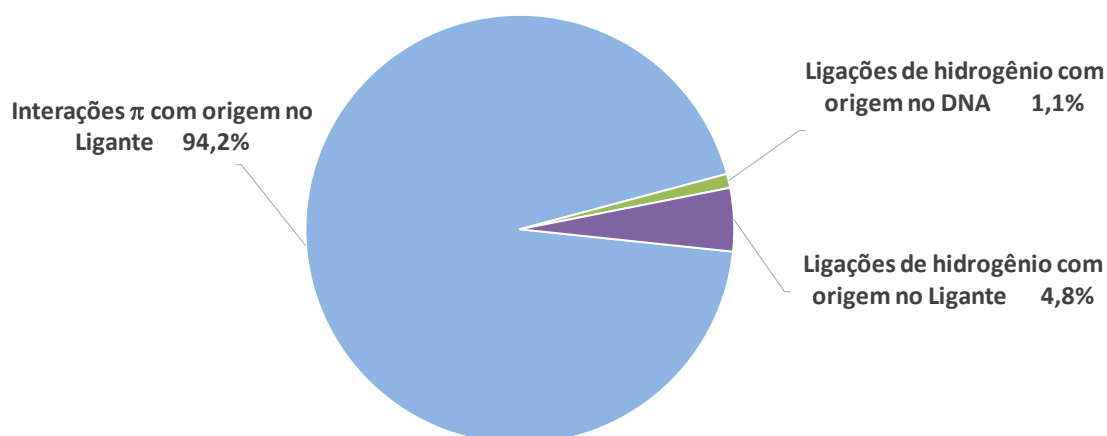


**Tabela 3.37** Estruturas de DNA com os ligantes dotados de anéis aromáticos localizados no sulco menor e interações mapeadas, tabela elaborada pelo autor.

Cód.	Interações $\pi$				Ligação de Hidrogênio				Relação Interação $\pi$ x Ligação de Hidrogênio
	Totais	Origem no DNA	Origem no ligante	% no ligante	Totais	Origem no DNA	Origem no ligante	% no ligante	
LS 1	15	0	15	100,00%	10	2	8	80,00%	60,00%
LS 2	7	0	7	100,00%	1	0	1	100,00%	87,50%
LS 3	5	0	5	100,00%	1	0	1	100,00%	83,33%
LS 4	16	0	16	100,00%	0	0	0	0,00%	100,00%
LS 5	16	0	16	100,00%	2	0	2	100,00%	88,89%
LS 6	4	0	4	100,00%	0	0	0	0,00%	100,00%
LS 7	10	0	10	100,00%	2	0	2	100,00%	83,33%
LS 8	11	0	11	100,00%	2	2	0	0,00%	84,62%
LS 9	15	0	15	100,00%	1	0	1	100,00%	93,75%
LS 10	10	0	10	100,00%	0	0	0	0,00%	100,00%
LS 11	10	0	10	100,00%	2	1	1	50,00%	83,33%
LS 12	742	0	742	100,00%	2	0	2	100,00%	99,73%
LS 13	9	0	9	100,00%	2	0	2	100,00%	81,82%
LS 14	14	0	14	100,00%	1	0	1	100,00%	93,33%
LS 15	19	0	19	100,00%	1	0	1	100,00%	95,00%
LS 16	17	0	17	100,00%	0	0	0	0,00%	100,00%
LS 17	15	0	15	100,00%	1	0	1	100,00%	93,75%
LS 18	6	0	6	100,00%	4	1	3	75,00%	60,00%
LS 19	5	0	5	100,00%	0	0	0	0,00%	100,00%
LS 20	15	0	15	100,00%	1	0	1	100,00%	93,75%
LS 21	1	0	1	100,00%	0	0	0	0,00%	100,00%
LS 22	9	0	9	100,00%	0	0	0	0,00%	100,00%
LS 23	6	0	6	100,00%	0	0	0	0,00%	100,00%
LS 24	9	0	9	100,00%	2	0	2	100,00%	81,82%
LS 25	8	0	8	100,00%	2	0	2	100,00%	80,00%
LS 26	6	0	6	100,00%	1	1	0	0,00%	85,71%
LS 27	8	0	8	100,00%	0	0	0	0,00%	100,00%
LS 28	9	0	9	100,00%	3	0	3	100,00%	75,00%
LS 29	9	0	9	100,00%	1	0	1	100,00%	90,00%
LS 30	20	0	20	100,00%	4	2	2	50,00%	83,33%
LS 31	9	0	9	100,00%	4	1	3	75,00%	69,23%
LS 32	17	0	17	100,00%	4	0	4	100,00%	80,95%
LS 33	6	0	6	100,00%	0	0	0	0,00%	100,00%
LS 34	6	0	6	100,00%	3	2	1	33,33%	66,67%
LS 35	4	0	4	100,00%	0	0	0	0,00%	100,00%
LS 36	12	0	12	100,00%	2	0	2	100,00%	85,71%
LS 37	10	0	10	100,00%	1	1	0	0,00%	90,91%
LS 38	17	0	17	100,00%	0	0	0	0,00%	100,00%
LS 39	132	0	132	100,00%	1	0	1	100,00%	99,25%
LS 40	15	0	15	100,00%	2	0	2	100,00%	88,24%
LS 41	20	0	20	100,00%	2	0	2	100,00%	90,91%

<i>Continuação da Tabela 3.37</i>									
LS 42	12	0	12	100,00%	1	0	1	100,00%	92,31%
LS 43	17	0	17	100,00%	1	0	1	100,00%	94,44%
LS 44	13	0	13	100,00%	2	0	2	100,00%	86,67%
LS 45	16	0	16	100,00%	2	0	2	100,00%	88,89%
LS 46	7	0	7	100,00%	1	0	1	100,00%	87,50%
LS 47	6	0	6	100,00%	1	0	1	100,00%	85,71%
LS 52	7	0	7	100,00%	6	2	4	66,67%	53,85%
LS 53	10	0	10	100,00%	2	0	2	100,00%	83,33%
LS 54	9	0	9	100,00%	2	0	2	100,00%	81,82%
LS 55	11	0	11	100,00%	2	1	1	50,00%	84,62%
LS 56	16	0	16	100,00%	1	0	1	100,00%	94,12%
LS 57	4	0	4	100,00%	2	0	2	100,00%	66,67%
Total geral	1422	0	1422	100,00%	88	16	72	81,82%	94,17%

Constata-se pelos dados da Tabela 3.37 que 100% das interações  $\pi$  existentes entre o DNA e o ligante no sulco possuem origem nos centroides do ligante, enquanto que a maioria das ligações de hidrogênio (81,82%) também apresentam a mesma origem. Pode-se ainda afirmar que quase a totalidade das interações DNA – ligante é efetuada pelas interações  $\pi$ , sendo dessa forma respondida a segunda questão, referente à forma como os ligantes interagem com a estrutura do DNA. A Figura 3.8 exibe graficamente a proporção dessas ligações por tipo.



**Figura 3.8** Distribuição dos tipos de ligações entre o DNA e os seus ligantes com anéis aromáticos no sulco, gráfico elaborado pelo autor.

Os dados acima relacionados vêm a trazer uma nova visão sobre o papel das interações  $\pi$  nos complexos ligante-DNA existentes no sulco menor, pois conforme Morávek e colegas (Morávek et al., 2002), até então estas ligações eram efetuadas principalmente pelas ligações de hidrogênio e forças de van der Waals, conceito este também compartilhado por diversos outros autores (Pullman, 1985; Davis e Teague, 1999; Brooijmans e Kuntz, 2003; Shaikh e Jayaram, 2012).

### 3.5 RELAÇÃO ENTRE INTERAÇÕES $\pi$ E LIGAÇÕES DE HIDROGÊNIO NA ESTRUTURA DE DNA

De acordo com as abordagens anteriormente efetuadas, a estrutura do DNA é resultado da atuação das clássicas ligações de hidrogênio entre os seus pares de bases (Watson e Crick, 1953; Crick e Watson, 1954; Quigley et al., 1986; Fonseca Guerra et al., 2000) e também da atuação das interações  $\pi$  entre as bases de uma mesma fita, promovendo o seu empilhamento (Mignon et al., 2005; Ermondi e Caron, 2006; Tewari e Dubey, 2008; Seng et al., 2010). Todavia, após o mapeamento vetorial efetuado pelo WIM e tendo o conhecimento das interações existentes entre bases de fitas diferentes e envolvendo as guaninas DG2, DG4, DG10, DG12, DG14, DG16, DG22 e DG24, e sabendo que praticamente a totalidade dos ligantes aromáticos interagem com o sulco menor das estruturas de DNA via interações  $\pi$ , surgiu uma terceira e última questão: haveria uma relação entre a quantidade de interações  $\pi$  entre fitas e a quantidade de ligações de hidrogênio, também entre estas fitas? Para que seja obtida essa resposta, foram analisadas as interações não covalentes em grupos de estruturas cristalográficas de DNA organizados pela presença ou não de ligantes, bem como o seu tipo, caso existam. Os resultados são mostrados nas tabelas 3.38 a 3.42.

**Tabela 3.38** Relações das interações não covalentes mapeadas pelo WIM para a estrutura de DNA com o ligante intercalado, tabela elaborada pelo autor

Cód.	Interações $\pi$				H Bond				Relação H Bond x interações $\pi$ (entre fitas)
	Totais	Na mesma fita	Entre fitas	% entre fitas	Totais	Na mesma fita	Entre fitas	% entre fitas	
LI 1	167	158	9	5,4%	28	1	27	96,4%	3,00

**Tabela 3.39** Relações das interações não covalentes mapeadas pelo WIM para as estruturas de DNA sem ligantes, tabela elaborada pelo autor

Cód.	Interações $\pi$				H Bond				Relação H Bond x interações $\pi$ (entre fitas)
	Totais	Na mesma fita	Entre fitas	% entre fitas	Totais	Na mesma fita	Entre fitas	% entre fitas	
SL 1	145	118	27	18,6%	42	7	35	83,3%	1,30
SL 2	194	182	12	6,2%	32	2	30	93,8%	2,50
SL 3	128	99	29	22,7%	42	11	31	73,8%	1,07
SL 4	198	182	16	8,1%	33	2	31	93,9%	1,94
SL 5	125	110	15	12,0%	34	2	32	94,1%	2,13
SL 6	174	159	15	8,6%	31	0	31	100,0%	2,07
SL 7	132	111	21	15,9%	33	3	30	90,9%	1,43
SL 8	181	167	14	7,7%	32	2	30	93,8%	2,14
SL 9	132	102	30	22,7%	32	0	32	100,0%	1,07
SL 10	182	165	17	9,3%	32	0	32	100,0%	1,88
SL 11	149	141	8	5,4%	38	8	30	78,9%	3,75
SL 12	129	126	3	2,3%	26	0	26	100,0%	8,67
SL 13	151	138	13	8,6%	32	0	32	100,0%	2,46
SL 14	199	181	18	9,0%	32	1	31	96,9%	1,72
SL 15	152	134	18	11,8%	36	2	34	94,4%	1,89
SL 16	115	99	16	13,9%	28	0	28	100,0%	1,75

**Tabela 3.40** Relações das interações não covalentes mapeadas pelo WIM para as estruturas de DNA íons aprisionados por cadeia de moléculas de H<sub>2</sub>O no sulco menor, tabela elaborada pelo autor

Cód.	Interações $\pi$				H Bond				Relação H Bond x interações $\pi$ (entre fitas)
	Totais	Na mesma fita	Entre fitas	% entre fitas	Totais	Na mesma fita	Entre fitas	% entre fitas	
LS 48	139	124	15	10,8%	34	7	27	79,4%	1,80
LS 49	177	161	16	9,0%	32	1	31	96,9%	1,94
LS 50	168	153	15	8,9%	34	1	33	97,1%	2,20
LS 51	178	163	15	8,4%	31	2	29	93,5%	1,93
LS 60	186	171	15	8,1%	36	2	34	94,4%	2,27

**Tabela 3.41** Relações das interações não covalentes mapeadas pelo WIM para as estruturas de DNA com os ligantes sem anéis aromáticos no sulco menor, tabela elaborada pelo autor

Cód.	Interações $\pi$				H Bond				Relação H Bond x interações $\pi$ (entre fitas)
	Totais	Na mesma fita	Entre fitas	% entre fitas	Totais	Na mesma fita	Entre fitas	% entre fitas	
LS 58	146	135	11	7,5%	38	7	31	81,6%	2,82
LS 59	142	131	11	7,7%	38	7	31	81,6%	2,82

**Tabela 3.42** Relações das interações não covalentes mapeadas pelo WIM para as estruturas de DNA com os ligantes dotados de anéis aromáticos no sulco menor, tabela elaborada pelo autor.

Cód.	Interações $\pi$				H Bond				Relação H Bond x interações $\pi$ (entre fitas)
	Totais	Na mesma fita	Entre fitas	% entre fitas	Totais	Na mesma fita	Entre fitas	% entre fitas	
LS 1	153	142	11	7,2%	37	5	32	86,5%	2,91
LS 2	144	129	15	10,4%	32	0	32	100,0%	2,13
LS 3	165	156	9	5,5%	32	1	31	96,9%	3,44
LS 4	169	157	12	7,1%	32	0	32	100,0%	2,67
LS 5	158	146	12	7,6%	38	36	33	86,8%	2,75
LS 6	180	167	13	7,2%	33	3	30	90,9%	2,31
LS 7	163	149	14	8,6%	34	2	32	94,1%	2,29
LS 8	134	124	10	7,5%	36	1	35	97,2%	3,50
LS 9	136	124	12	8,8%	32	0	32	100,0%	2,67
LS 10	136	122	14	10,3%	41	7	34	82,9%	2,43
LS 11	136	126	10	7,4%	29	3	26	89,7%	2,60
LS 12	180	171	9	5,0%	24	3	21	87,5%	2,33
LS 13	135	123	12	8,9%	31	4	27	87,1%	2,25
LS 14	185	171	14	7,6%	34	2	32	94,1%	2,29
LS 15	151	139	12	7,9%	34	4	30	88,2%	2,50
LS 16	179	166	13	7,3%	31	1	30	96,8%	2,31
LS 17	152	133	19	12,5%	29	0	29	100,0%	1,53
LS 18	149	135	14	9,4%	33	1	32	97,0%	2,29
LS 19	143	131	12	8,4%	32	0	32	100,0%	2,67
LS 20	136	124	12	8,8%	32	0	32	100,0%	2,67
LS 21	145	136	9	6,2%	37	5	32	86,5%	3,56
LS 22	181	167	14	7,7%	30	0	30	100,0%	2,14
LS 23	193	180	13	6,7%	33	4	29	87,9%	2,23
LS 24	186	172	14	7,5%	34	4	30	88,2%	2,14
LS 25	128	115	13	10,2%	33	1	32	97,0%	2,46
LS 26	191	182	9	4,7%	34	3	31	91,2%	3,44
LS 27	182	168	14	7,7%	33	4	29	87,9%	2,07
LS 28	135	124	11	8,1%	36	6	30	83,3%	2,73
LS 29	127	119	8	6,3%	36	4	32	88,9%	4,00
LS 30	136	125	11	8,1%	38	5	33	86,8%	3,00
LS 31	177	165	12	6,8%	31	1	30	96,8%	2,50
LS 32	176	164	12	6,8%	33	1	32	97,0%	2,67
LS 33	130	112	18	13,8%	31	1	30	96,8%	1,67
LS 34	189	174	15	7,9%	33	4	29	87,9%	1,93
LS 35	180	167	13	7,2%	33	3	30	90,9%	2,31
LS 36	171	155	16	9,4%	36	4	32	88,9%	2,00
LS 37	173	159	14	8,1%	31	0	31	100,0%	2,21
LS 38	95	91	4	4,2%	29	0	29	100,0%	7,25
LS 39	169	157	12	7,1%	35	2	33	94,3%	2,75

*Continuação da Tabela 3.42*

LS 40	168	154	14	8,3%	31	0	31	100,0%	2,21
LS 41	166	152	14	8,4%	31	0	31	100,0%	2,21
LS 42	120	106	14	11,7%	34	0	34	100,0%	2,43
LS 43	169	155	14	8,3%	32	0	32	100,0%	2,29
LS 44	136	123	13	9,6%	33	1	32	97,0%	2,46
LS 45	171	162	9	5,3%	35	2	33	94,3%	3,67
LS 46	164	151	13	7,9%	32	0	32	100,0%	2,46
LS 47	181	166	15	8,3%	30	1	29	96,7%	1,93
LS 52	123	105	18	14,6%	38	0	38	100,0%	2,11
LS 53	106	96	10	9,4%	37	1	36	97,3%	3,60
LS 54	170	155	15	8,8%	33	1	32	97,0%	2,13
LS 55	181	165	16	8,8%	33	1	32	97,0%	2,00
LS 56	172	159	13	7,6%	32	0	32	100,0%	2,46
LS 57	146	132	14	9,6%	37	4	33	89,2%	2,36

As tabelas 3.38 a 3.42 permitem que sejam corroborados os conceitos relativos à manutenção da dupla hélice pelas ligações de hidrogênio pelos pares de bases A-T e G-C, ao papel das interações  $\pi$  no empilhamento das bases, formando a hélice. Como novos conceitos, identificou-se a existência de uma relação entre a quantidade de interações  $\pi$  entre fitas e a quantidade de ligações de hidrogênio, também entre estas fitas. A Tabela 3.42 mostra uma análise do comportamento das estruturas do grupo de estruturas de DNA com a sequência de nucleotídeos (\*CP\*GP\*CP\*GP\*AP\*AP\*TP\*TP\*CP\*GP\*CP\*G), em função das suas interações entre fitas.

**Tabela 3.43** Relações das interações não covalentes mapeadas pelo WIM para as estruturas de DNA, tabela elaborada pelo autor.

Característica da estrutura do DNA em função dos ligantes	Interações $\pi$ hidrogênio entre fitas		Ligações de hidrogênio entre fitas		Relação H Bond x interações $\pi$ (entre fitas)	
	Média	Desvio Padrão	Média	Desvio Padrão	Média	Desvio Padrão
SL - Sem ligantes	17,0	2,6	30,9	2,0	2,4	1,7
LS - Ligantes com anéis aromáticos	12,7	6,9	31,2	2,5	2,6	0,8
LS - Íons no sulco menor	15,2	0,4	30,8	2,6	2,0	0,2
LS - Ligantes sem anéis aromáticos	11,0	0,0	31,0	0,0	2,8	0,0

A tabela 3.43 nos fornece uma ideia de como a presença de ligantes com anéis aromáticos interfere na quantidade de interações  $\pi$  entre fitas, fato este que não afeta consideravelmente as ligações de hidrogênio entre os pares de bases A-T e G-C. Outro fato a ser constatado, é que indiferentemente da presença de ligante, do seu tipo ou da ausência destes, a proporção existente entre as interações  $\pi$  entre fitas e as ligações de hidrogênio entre os pares de bases A-T e G-C apresenta-se dentro de um padrão, excluindo-se as estruturas com íons aprisionados no sulco menor.

# Capítulo 4

---

CONCLUSÕES



#### 4. MAPEAMENTO VETORIAL

O modelo matemático concebido inicialmente para o mapeamento das interações  $\pi$ , utilizando o cálculo vetorial aplicado às coordenadas cristalográficas dos átomos de biomoléculas (proteínas e DNA) permitiu caracterizar as interações  $\pi$  ( $\text{CH}\cdots\pi$ ,  $\text{NH}\cdots\pi$ ,  $\pi\cdots\pi$  e outras), bem como obter uma nova visão sobre a ação desse tipo de interações não covalentes na manutenção da estrutura do DNA contribuindo para o empilhamento de suas bases. Em função dos resultados obtidos e considerando que as ligações de hidrogênio também são regidas por parâmetros geométricos específicos e também pela utilidade do conceito abordado para a química supramolecular, justificou-se a transformação do algoritmo utilizado em uma ferramenta computacional a qual foi denominada WIM (*Weak Interaction Mapping*), que transformou o mapeamento vetorial em ferramenta de pesquisa.

Paralelamente a isso, em trabalhos anteriores constatou-se que além das interações não-covalentes, tais como as interações de van der Waals e ligações de hidrogênio, foram encontrados indícios de que as interações  $\pi$  também pudessem contribuir para a existência do formato de dupla hélice do DNA, uma vez que foram encontradas a existência destas entre bases pertencentes a fitas diferentes. Em função disso, este trabalho efetuou um estudo comparativo entre as ligações de hidrogênio e interações  $\pi$  obtidas por mapeamento vetorial em estruturas de DNA selecionadas em função da quantidade de pares de bases e a sequência das fitas (com e sem ligante).

Utilizando o WIM como ferramenta de pesquisa, foi possível mapear todas as interações  $\pi$  e ligações de hidrogênio existentes em 77 estruturas de DNA com ou sem ligantes e com a sequência de nucleotídeos (\*CP\*GP\*CP\*GP\*AP\*AP\*TP\*TP\*CP\*GP\*CP\*G) e obter para estas, arquivos compatíveis para uso tanto em planilhas eletrônicas quanto visualizadores gráficos. O output desse processamento forneceu dados suficientes para

que fossem efetuadas análises sobre o comportamento destas estruturas em função das interações  $\pi$  e ligações de hidrogênio, bem como a presença ou não de ligantes.

O primeiro dos resultados obtido foi a descoberta de que existe um padrão na ocorrência de interações  $\pi$  entre fitas, sendo que estas sempre possuem a sua origem no centroide de do anel de 6 átomos de uma guanina (Ct6) e tem como destinos específicos os átomos mapeados H1, H21, H22, N1, N2 e O6, além do Ct6 das guaninas DG2, DG4, DG10, DG12, DG14, DG16, DG22 2 DG24. Com exceção de uma única estrutura de DNA que possui o ligante intercalado induzindo uma deformação que reduziu o número de interações  $\pi$ ; nas demais estruturas espaciais com ligantes no sulco ou sem as quantidades destas interações o comportamento se apresenta semelhante para todas elas. Tal padrão é justificado pelo direcionamento das bases para o eixo da dupla hélice e isso faz com que os centroides do anel de 06 átomos das guaninas especificadas e os átomos H1, H21, H22, N1, N2 e O6, além do Ct6 possuam coordenadas espaciais propícias à ocorrência das interações  $\pi$ .

Outro resultado alcançado foi que dentre o grupo de estruturas selecionadas, em análises efetuadas em um grupo de 53 estruturas que possuem no seu sulco menor ligantes com anéis aromáticos, foi que as interações DNA – ligante são efetuadas majoritariamente pelas interações  $\pi$  com origem nos ligantes e estas correspondem a 94,2% do total, as ligações de hidrogênio com origem nos ligantes 4,8% e por fim, as ligações de hidrogênio com origem no DNA com apenas 1,1%. Um outro fato constatado foi que 100% das interações  $\pi$  existentes entre o DNA e o ligante no sulco possuem origem nos centroides dos ligantes, enquanto que a maioria das ligações de hidrogênio (81,82%) também apresentam a mesma origem. Os dados acima relacionados vêm a trazer uma nova visão sobre o papel das interações  $\pi$  nos complexos ligante-DNA existentes no sulco menor, pois até então, concebia-se que estas ligações eram efetuadas principalmente pelas ligações de hidrogênio e forças de van der Waals. Dessa forma, pode-se concluir que um complexo ou fármaco deve possuir anéis aromáticos para uma melhor ligação mais eficiente no sulco menor do DNA.

De acordo a literatura pesquisada, a estrutura do DNA é o resultado da atuação das clássicas ligações de hidrogênio entre os pares de bases A-T e G-C, e também pela atuação das interações  $\pi$  entre as bases de uma mesma fita, promovendo o seu empilhamento, fatos estes corroborados pelo mapeamento vetorial. Todavia, após a descoberta das interações existentes entre bases de fitas diferentes e envolvendo as guaninas DG2, DG4, DG10, DG12, DG14, DG16, DG22 2 DG24, e sabendo que praticamente a totalidade dos ligantes aromáticos interagem com o sulco menor das estruturas de DNA via interações  $\pi$ , buscou-se a resposta à uma última questão, a qual seria: se existe um padrão para a ocorrência das interações entre fitas do DNA, na quantidade existente das ligações de hidrogênio entre os pares de bases A-T e G-C e ainda uma relação proporcional entre estas duas interações não covalentes. Como consequência das análises efetuadas, identificamos a existência de uma relação entre a quantidade de interações  $\pi$  entre fitas e a quantidade de ligações de hidrogênio, também entre estas mesmas fitas. Obteve-se uma nova ideia de como a presença de ligantes com anéis aromáticos interfere na quantidade de interações  $\pi$  entre fitas, fato este que não afeta consideravelmente as ligações de hidrogênio entre os pares de bases A-T e G-C. Outro fato constatado, é que indiferentemente da presença ou não de ligantes, a proporção existente entre as interações  $\pi$  entre fitas e as ligações de hidrogênio entre os pares de bases A-T e G-C apresenta-se dentro de um padrão, excluindo-se as estruturas com íons aprisionados no sulco menor.

Com os resultados obtidos nessa pesquisa, conclui-se que a automatização do mapeamento vetorial obtida com a criação do *software* WIM para a pesquisa das interações supramoleculares envolvendo anéis aromáticos e ligações de hidrogênio, permitiu obter novos conhecimentos sobre a importância destas interações não covalentes na estrutura espacial do DNA. São novas informações sobre como estas interações ocorrem nos complexos ligante-DNA e também a constatação da existência de uma relação proporcional entre a quantidades das interações  $\pi$  entre fitas e das ligações de hidrogênio nos pares de bases A-T e G-C. Esse novo conhecimento vem a consolidar o WIM como ferramenta auxiliar para o estudo de biomoléculas (proteínas e DNA) em conjunto com

softwares para visualização gráfica e dinâmica e *docking* moleculares. Por fim, este trabalho vem a preencher uma lacuna existente até este momento no campo da pesquisa de uma das interações eletrostáticas importantes em macromoléculas.

Outro aspecto a ser considerado refere-se ao ganho de produtividade na pesquisa obtido com o WIM. Sem os recursos disponibilizados pelo software, seriam necessários anos de trabalho e análise para que os mesmos resultados apresentados para as 77 estruturas de DNA fossem alcançados.

Considerando a funcionalidade do software, o WIM possui além de desafios a serem vencidos, oportunidades a serem conquistadas, conforme relacionado a seguir.

#### Desafios:

- Elaborar um algoritmo para a inclusão na estrutura os átomos de hidrogênio para os arquivos baixados diretamente da nuvem (PDB-RSCB);
- Otimizar o tempo do processamento;
- Expandir a metodologia utilizada para outras sequências de DNA.

#### Oportunidades:

- Na Dinâmica Molecular, explorar a possibilidade de quantificar as interações em frames específicos e mensurar qualitativa e quantitativamente as interações existentes;
- Para o Docking Molecular considerar a possibilidade de análise das interações resultantes do docking;
- Aplicação em biossensores para análise *in silico* de biossensores e suas interações;
- Na pesquisa de fármacos, avaliar as interações do fármaco e complexos resultantes.

# Referências

Aakeroy, C. B.; Seddon, K. R. The hydrogen bond and crystal engineering. *Chemical Society Reviews* **1993**, v. 22, n. 6, p. 397-407

Accelrys Discovery Studio Visualizer™ - Accelrys Software Inc. Discovery Studio Visualizer Client 3.1

Adams, A. et al. Structures of two minor-groove-binding quinolinium quaternary salts complexed with d(CGCGAATTCGCG)(2) at 1.6 and 1.8 Angstrom resolution. *Acta Crystallographica, Section D*, **2005** v. 61, p. 1348-1353.

ARIS Express - Free Modeling Software. ARIS BPM Community: IDS Scheer 2009.

Babu, M. M.: NCI: a server to identify non-canonical interactions in protein structures. *Nucleic Acids Research* **2003**, 31, 3345-3348.

Bao, G. Mechanics of biomolecules. *Journal of the Mechanics and Physics of Solids* **2002**, v. 50, n. 11, p. 2237-2274.

Berman, H. M. et al. The Nucleic Acid Database (NDB). In: Rossmann, M. G. e Arnold, E. (Ed.). *International Tables for Crystallography Volume F: Crystallography of biological macromolecules*. Dordrecht: Springer Netherlands, **2001**. p.657-662.

Berman, H. M.; Young, P. R. The Interaction of Intercalating Drugs with Nucleic Acids. *Annual Review of Biophysics and Bioengineering* 1981, v. 10, n. 1, p. 87-114.

Berman, H. M.; Henrick, K.; Nakamura, H.; Markley, J. L.: The worldwide Protein Data Bank (wwPDB): ensuring a single, uniform archive of PDB data. *Nucleic Acids Research* **2007**, 35, D301-D303.

Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, p. E.: The Protein Data Bank. *Nucleic Acids Research* **2000**, 28, 235-242.

Berman, H. M.; Young, P. R. The Interaction of Intercalating Drugs with Nucleic Acids. *Annual Review of Biophysics and Bioengineering* **1981**, v. 10, n. 1, p. 87-114.

BioMat - Grupo de Física Teórica de Materiais e Biomoléculas.  
<http://www.df.ufscar.br/biomat-grupo-de-fisica-teorica-de-materiais-e-biomoleculas>,  
acesso em 15-08-2016.

Birchall, L. S.; Roy, S.; Jayawarna, V.; Hughes, M.; Irvine, E.; Okorogheye, G. T.; Saudi, N.; De Santis, E.; Tuttle, T.; Edwards, A. A.; Ulijn, R. V.: Exploiting CH- $\pi$  interactions in supramolecular hydrogels of aromatic carbohydrate amphiphiles. *Chemical Science* **2011**, 2, 1349-1355.

Brooijmans, N.; Kuntz, I. D. Molecular recognition and docking algorithms. *Annual Review of Biophysics and Biomolecular Structure*, **2003** v. 32, n. 1, p. 335-373.

Brown, D. G. et al. Crystal structure of a berenil-dodecanucleotide complex: the role of water in sequence-specific ligand binding. *The EMBO Journal* **1990**, v. 9, p. 1329-1334.

Campbell, N. H. et al. Targeting the DNA minor groove with fused ring dicationic compounds: Comparison of in silico screening and a high-resolution crystal structure. *Bioorganic & Medicinal Chemistry Letters* **2006**, v. 16, n. 1, p. 15-19.

Caracelli, I.; Zukerman-Schpector, J.; Sacco, A. C. S.; Pereira, F.: A informática na pesquisa em biomoléculas. VIII Simpósio de Ciência e Tecnologia da FATEC Tatuí, **2013**, Tatuí. p.55-57.

Caracelli, I.; Zukerman-Schpector, J.; Sacco, A. C. S.; Pereira, F.: Mapeamento vetorial de interações  $\pi$  em macromoléculas: busca por padrões em grupo específico de DNA. IX Simpósio de Ciência e Tecnologia da FATEC Tatuí, **2014**, Tatuí. FATEC Tatuí. p.75-76.

Caracelli, I.; Zukerman-Schpector, J.; Sacco, A. C. S.; Souza, L. M.: Requisitos de um projeto de software para o cálculo de interações  $\pi$ . X Simpósio de Ciência e Tecnologia da FATEC

Tatuí, **2015**, Tatuí. FATEC Tatuí. p.145-146.

Caracelli, I.; Zukerman-Schpector, J.; Sacco, A. C. S.; Souza, L. M.: Desenvolvimento de software para cálculo de interações  $\pi$ . Revista Sapere. Tatuí: FATEC Tatuí. 8: 1-5 p. **2016**.

Caracelli, I.; Zukerman-Schpector, J.; Araujo-Moreira, F. M. BioMat - Grupo de Física Teórica de Materiais e Biomoléculas - Departamento de Física - UFSCar. **2012**.

Caracelli, I.; Zukerman-Schpector, J.; Sacco, A. C. S.; Pereira, F.: Mapeamento vetorial de interações  $\pi$  em macromoléculas: busca por padrões em grupo específico de DNA. Revista Sapere. Tatuí: FATEC Tatuí. 6: 1-13 p. **2014**.

Caracelli, I.; Zukerman-Schpector, J.; Maganhi, S. H.; Stefani, H. A.; Guadagnin, R.; Tiekink, E. R. T.: 2-chlorovinyl tellurium dihalides, (p-tol)Te[C(H)=C(Cl)Ph]X<sub>2</sub> for X = Cl, Br and I: variable coordination environments, supramolecular structures and docking studies in cathepsin B. *Journal of the Brazilian Chemical Society* **2010**, 21, 2155-2163.

Caracelli, I.; Zukerman-Schpector, J.: PPG BIOTEC BIT 603 - Biomoléculas. São Carlos: UFSCar **2009**.

Caracelli, I.; Zukerman-Schpector, J.: PPG BIOTEC BIT 603 - Fundamentos da Estrutura de Proteínas - Parte 1. São Carlos: UFSCar **2010**.

Caracelli, I.; Zukerman-Schpector, J.; Tiekink, E. R. T.: Supramolecular aggregation patterns based on the bio-inspired Se(lone pair)··· $\pi$ (aryl) synthon. *Coordination Chemistry Reviews* **2012**, 256, 412-438.

Cauët, E. et al. Histidine–Aromatic Interactions in Proteins and Protein–Ligand Complexes: Quantum Chemical Study of X-ray and Model Structures. *Journal of Chemical Theory and Computation* **2005**, v. 1, n. 3, p. 472-483.



Cerny, J.; Hobza, P. Non-covalent interactions in biomacromolecules. *Physical Chemistry Chemical Physics* **2007**, v. 9, n. 39, p. 5291-5303.

Chatake, T. et al. Crystallographic studies on damaged DNAs. II. N(6)-methoxyadenine can present two alternate faces for Watson-Crick base-pairing, leading to pyrimidine transition mutagenesis. *Journal of Molecular Biology* **1999**, v. 294, p. 1223-1230.

Clark, G. R. et al. Structure of a bis-amidinium derivative of hoechst 33258 complexed to dodecanucleotide d(CGCGAATTCGCG)<sub>2</sub>: the role of hydrogen bonding in minor groove drug-DNA recognition. *Nucleic Acids Research* **1997**, v. 25, p. 1510-1515.

\_\_\_\_\_. Designer DNA-binding drugs: the crystal structure of a meta-hydroxy analogue of Hoechst 33258 bound to d(CGCGAATTCGCG)<sub>2</sub>. *Nucleic Acids Research* **1996**, v. 24, p. 4882-4889.

Clausen-Schaumann, H. et al. Mechanical Stability of Single DNA Molecules. *Biophysical Journal* **2000**, v. 78, n. 4, p. 1997-2007.

Crick, F. H. C.; Watson, J. D. The Complementary Structure of Deoxyribonucleic Acid. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **1954**, v. 223, n. 1152, p. 80-96.

Cunha, R. L. O. R.; Zukerman-Schpector, J.; Caracelli, I.; Comasseto, J.: Revisiting the addition reaction of TeCl<sub>4</sub> to alkynes: The crystal structure and docking studies of 1-chloro-2-trichlorotelluro-3-phenyl-propen-2-ol. *Journal of Organometallic Chemistry* **2006**, 691, 4807-4815.

DAVIS, A. M.; TEAGUE, S. J. Hydrogen bonding, hydrophobic interactions, and failure of the rigid receptor hypothesis. *Angewandte Chemie International Edition* **1999**, v. 38, n. 6, p. 736-749.

Denisov, A. Y. et al. The solution conformation of a carbocyclic analog of the Dickerson-Drew dodecamer: comparison with its own X-ray structure and that of the NMR structure of the native counterpart. *Journal of Biomolecular Structure and Dynamics* **1998**, v. 16, p.

547-568.

Dougherty, D. A.: Cation- $\pi$  Interactions Involving Aromatic Amino Acids. *The Journal of Nutrition* **2007**, 137, 1504S-1508S.

Drew, H. R.; Samson, S.; Dickerson, R. E. Structure of a B-DNA dodecamer at 16 K. *Proceedings of the National Academy of Sciences* **1982**, v. 79, p. 4040-4044.

Drew, H. R.; Wing, R. M.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E.: Structure of a B-DNA dodecamer: conformation and dynamics. *Proceedings of the National Academy of Sciences* **1981**, 78, 2179-2183.

Du, Q.-S. et al. Empirical formulation and parameterization of cation- $\pi$  interactions for protein modeling. *Journal of Computational Chemistry* **2012**, v. 33, n. 2, p. 153-162.

\_\_\_\_\_. Energies and physicochemical properties of cation- $\pi$  interactions in biological structures. *Journal of Molecular Graphics and Modelling* **2012**, v. 34, n. 0, p. 38-45

Edwards, K. J.; Jenkins, T. C.; Neidle, S. Crystal structure of a pentamidine-oligonucleotide complex: implications for DNA-binding properties. *Biochemistry* 1992, v. 31, p. 7104-7109.

EMBL-EBI, Protein Data Bank Summary (PDBsum): <http://www.ebi.ac.uk/thornton-srv/databases/pdbsum/>, acesso em 05-04-2012.

Ermondi, G.; Caron, G.: Recognition forces in ligand-protein complexes: Blending information from different sources. *Biochemical Pharmacology* **2006**, 72, 1633-1645.

FAI. Agência Inovação da UFSCar. São Carlos, 2017. <http://www.inovacao.ufscar.br/>, acesso em 10-07-2017.

Ferguson, L. R.; Denny, W. A. Genotoxicity of non-covalent interactions: DNA intercalators. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 2007, v.

623.

Fonseca Guerra, C. et al. Hydrogen Bonding in DNA Base Pairs: Reconciliation of Theory and Experiment. *Journal of the American Chemical Society* **2000**, v. 122, n. 17, p. 4117-4128.

Fujii, S. et al. Sequence-dependent DNA deformability studied using molecular dynamics simulations. *Nucleic Acids Research* **2007**, v. 35, n. 18, p. 6063-6074.

Gallivan, J. P.; Dougherty, D. A.: Cation- $\pi$  interactions in structural biology. *Proceedings of the National Academy of Sciences* **1999**, v. 96, n. 17, p. 9459-9464.

Gao, Q. et al. Drug-induced DNA repair: X-ray structure of a DNA-ditercalinium complex. *Proceedings of the National Academy of Sciences* **1991**, v. 88, n. 6, p. 2422-2426.

Goodsell, D. S. et al. Structure of a dicationic monoimidazole lexitropsin bound to DNA. *Biochemistry* **1995**, v. 34, p. 16654-16661.

Holbrook, S. R.; Dickerson, R. E.; Kim, S.-H. Anisotropic Thermal-Parameter Refinement of the DNA Dodecamer CGCGAATTCGCG by the Segmented Rigid-Body Method. *Acta Crystallographica, Section B* **1985**, v. 41, p. 255-262.

Howerton, S. B. et al. Locating monovalent cations in the grooves of B-DNA. *Biochemistry* **2001**, v. 40, n. 34, p. 10023-10031.

Hunter, C. A.; Lawson, K. R.; Perkins, J.; Urch, C. J., C. A.: Aromatic interactions. *Journal of the Chemical Society, Perkin Transactions 2* **2001**, 5, 651-669

Hunter, C. A.; Sanders, J. K. M. The nature of  $\pi$ - $\pi$  interactions. *Journal of the American Chemical Society* **1990**, 112, 14, 5525-5534.

Imai, Y. N. et al. Amide- $\pi$  interactions between formamide and benzene. *Journal of Computational Chemistry* **2009**, v. 30, n. 14, p. 2267-2276.

Janiak, C.: A critical account on [small pi]-[small pi] stacking in metal complexes with aromatic nitrogen-containing ligands. *Journal of the Chemical Society, Dalton Transactions* **2000**, 21, 3885-3896.

Jersey, T. S. U. O. N. Nucleic Acid Database (NDB). 2012. <http://ndbserver.rutgers.edu/> , acesso em 23-10-2012.

Karplus, P. A.; Schulz, G. E.: Substrate binding and catalysis by glutathione reductase as derived from refined enzyme: Substrate crystal structures at 2Å resolution. *Journal of Molecular Biology* **1989**, 210, 163-180.

Kielkopf, C. L. et al. A Structural Basis for Recognition of A·T and T·A Base Pairs in the Minor Groove of B-DNA. *Science* **1998**, v. 282, n. 5386, p. 111-115.

Kiviniemi, S.: *Complexation of N-heteroaromatic cations with crown ethers and tetraphenylborate*. **2001**. 70 p. Tese (Doutorado) - Faculty of Science, Department of Chemistry, University of Oulu, Finland.

Kollman, P. A.; Allen, L. C. Theory of the hydrogen bond. *Chemical Reviews* **1972**, v. 72, n. 3, p. 283-303.

Komeda, S. et al. A Third Mode of DNA Binding: Phosphate Clamps by a Polynuclear Platinum Complex. *Journal of the American Chemical Society* 2006, v. 128, p. 16092-16103.

Kool, E. T. Hydrogen bonding, base stacking, and steric effects in DNA replication. *Annual Review of Biophysics and Biomolecular Structure* **2001**, v. 30, n. 1, p. 1-22.

Kurumbail, R. G.; Stevens, A. M.; Gierse, J. K.; McDonald, J. J.; Stegeman, R. A.; Pak, J. Y.; Gildehaus, D.; Iyashiro, J. M.; Penning, T. D.; Seibert, K.; Isakson, p. C.; Stallings, W.: Structural basis for selective inhibition of cyclooxygenase-2 by anti-inflammatory agents. *Nature* **1996**, 384, 644-648.

Kuszewski, J.; Schwieters, C.; Clore, G. M. Improving the accuracy of NMR structures of DNA by means of a database potential of mean force describing base-base positional interactions. *Journal of the American Chemical Society* **2001**, v. 123, p. 3903-3918.

LaCrEMM - Laboratório de Cristalografia Estereodinâmica e Modelagem Molecular. <http://www.quimica.ufscar.br/>, acesso em 26-07-2011.

Larsen, T. A. et al. The structure of DAPI bound to DNA. *Journal of Biomolecular Structure and Dynamics* **1989**, v. 7, p. 477-491.

Laudon, J.; Laudon, K. C. Sistemas de informações gerenciais. São Paulo: Editora Pearson, 2004.

Laughton, C. A. et al. A crystallographic and spectroscopic study of the complex between d(CGCGAATTCGCG)<sub>2</sub> and 2,5-bis(4-guanylphenyl)furan, an analogue of berenil. Structural origins of enhanced DNA-binding affinity. *Biochemistry* 1996, v. 35, p. 5655-5661.

Lee, M. P. H.; Neidle, S. Crystal structure of the DB293-D(CGCGAATTCGCG)<sub>2</sub> complex. To be Published **2006a**.

\_\_\_\_\_. Crystal structure of the DB 911 - D(CGCGAATTCGCG)<sub>2</sub> complex at 2.05 Å resolution. To be Published **2006b**.

Lehn, J.-M.: Supramolecular chemistry: concepts and perspectives. Weinheim: VCH, c1995. 271 p.

Lercher, L. et al. Structural Insights Into How 5-Hydroxymethylation Influences Transcription Factor Binding. *Chemical Communications - Royal Society of Chemistry* **2014**, v. 50, p. 1794.

Levitt, M.; Perutz, M. F. Aromatic rings act as hydrogen bond acceptors. *Journal of Molecular Biology* **1988**, v. 201, n. 4, p. 751-754.

Lin, S.; Neidle, S.; Campbell, N. Crystal structure of the DB1880-D(CGCGAATTCGCG)<sub>2</sub> complex. To be Published **2010**.

Lindberg, D.; Caldwell, R.; Frankel, J. The finish line. Understanding Science, Berkeley, 2017. [http://undsci.berkeley.edu/article/0\\_0\\_0/dna\\_12](http://undsci.berkeley.edu/article/0_0_0/dna_12), acesso em 18-04-2017.

Liu, J. et al. The structure of the most studied DNA fragment changes under the influence of ions: a new packing of d(CGCGAATTCGCG). *FEBS Letters* **1998**, v. 438, p. 211-214.

Liu, J.; Subirana, J. A. Structure of d(CGCGAATTCGCG) in the presence of Ca (2+) ions. *Journal of Biological Chemistry* **1999**, v. 274, p. 24749-24752.

Ma, J. C.; Dougherty, D. A. The Cation- $\pi$  Interaction. *Chemical Reviews* **1997**, 97, 1303-1324.

Malinina, L.; Soler-Lopez, M.; Aymami, J.; Subirana, J. A.: Intercalation of an acridine-peptide drug in an AA/TT base step in the crystal structure of [d(CGCGAATTCGCG)]<sub>2</sub> with six duplexes and seven Mg(2+) ions in the asymmetric unit. *Biochemistry* **2002**, 41, 9341-9348.

Mallena, S.; Lee, M. P.; Bailly, C.; Neidle, S.; Kumar, A.; Boykin, D. W.; Wilson, W.: Thiophene-based diamidine forms a "super" at binding minor groove agent. *Journal of American Chemical Society* **2004**, 126, 13659-13669.

Mann, J. et al. A new class of symmetric bisbenzimidazole-based DNA minor groove-binding agents showing antitumor activity. *Journal of Medicinal Chemistry* **2001**, v. 44, p. 138-144.

Martin, R. F. et al. In Vitro Studies with Methylproamine. *Cancer Research* **2004**, v. 64, n. 3, p. 1067-1070.

\_\_\_\_\_. In vitro studies with methylproamine: a potent new radioprotector. *Cancer Research* **2004**, v. 64, p. 1067-1070.

Mazur, S. et al. A thermodynamic and structural analysis of DNA minor-groove complex formation. *Journal of Molecular Biology* **2000**, v. 300, n. 2, p. 321-337.

Mcgaughey, G. B.; Gagné, M.; Rappé, A. K.:  $\pi$ -Stacking Interactions. *Journal of Biological Chemistry* **1998**, 273, 15458-15463.

Menon, M. J.: Sobre as origens das definições dos produtos escalar e vetorial. *Revista Brasileira de Ensino de Física* **2009**, 31, 2305.1-2305.11.

Meyer, E. A.; Castellano, R. K.; Diederich, F.: Interactions with Aromatic Rings in Chemical and Biological Recognition. *Angewandte Chemie International Edition* **2003**, 42, 1210-1250.

Miao, Y. et al. Out-of-Shape DNA Minor Groove Binders: Induced Fit Interactions of Heterocyclic Dications with the DNA Minor Groove. *Biochemistry* **2005**, v. 44, p. 14701-14708.

Michael Gromiha, M.; Santhosh, C.; SuwA, M. Influence of cation- $\pi$  interactions in protein-DNA complexes. *Polymer* **2004**, v. 45, n. 2, p. 633-639.

Microsoft. Microsoft Corporation (**2010**). Microsoft Visual Studio 2010 Ultimate.

Microsoft, Microsoft Corporation (**2012**). Microsoft Office.

Microsoft, Microsoft Corporation (**2012**). Novidades no Microsoft Office Excel 2007.

Mignon, P.; Loverix, S.; Steyaert, J.; Geerlings, p.: Influence of the  $\pi$ - $\pi$  interaction on the hydrogen bonding capacity of stacked DNA/RNA bases. *Nucleic Acids Research* **2005**, 33, 1779-1789.

Mohan, N.; Vijayalakshmi, K. P.; Koga, N.; Suresh, C. H.: Comparison of aromatic NH... $\pi$ , OH... $\pi$ , and CH... $\pi$  interactions of alanine using MP2, CCSD, and DFT methods. *Journal of Computational Chemistry* **2010**, 31, 2874-2882.

Morávek, Z.; Neidle, S.; Schneider, B. Protein and drug interactions in the minor groove of DNA. *Nucleic Acids Research* **2002**, v. 30, n. 5, p. 1182-1191.

Munde, M. et al. Induced fit conformational changes of a "reversed amidine" heterocycle: optimized interactions in a DNA minor groove complex. *Journal of the American Chemical Society* **2007**, v. 129, p. 5688-5698, 2007

Munnur, D. G. et al. A Detailed Study of Water Network in the Minor Groove of D(CGCGAATTCGCG)<sub>2</sub> and D(CGCAAATTTGCG)<sub>2</sub> in Presence of Linear Groove Binders. To be Published **2012**.

Nechipurenko, Y; Grokhovsky, S; Gursky, G; Nechipurenko, D; Polozov, R.: DNA-Based Nanostructures: Changes of Mechanical Properties of DNA upon Ligand Binding. *NATO Science for Peace and Security Series* **2008**, 5, 59-67.

Neidle, S.; Lee, M. P. H. Crystal structure of a DB1055-D(CGCGAATTCGCG)<sub>2</sub> complex. To be Published **2006**.

Neidle, S. et al. Symmetric bis-benzimidazoles, a new class of sequence-selective DNA-binding molecules. *Journal of the Chemical Society, Chemical Communications* **1999**, p. 929-930.

Neidle, S.; Parkinson, G. N.; Lee, M. P. H. Crystal structure of the berenil-D(CGCGAATTCGCG)<sub>2</sub> complex at 1.65 Å resolution. To be Published **2006**.

Nelson, S. M.; Ferguson, L. R.; Denny, W. A. Non-covalent ligand/DNA interactions: Minor groove binding agents. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **2007**, v. 623, n. 1-2, p. 24-40.



Nguyen, B. et al. Strong Binding in the DNA Minor Groove by an Aromatic Diamidine With a Shape That Does Not Match the Curvature of the Groove. *Journal of the American Chemical Society* **2002**, v. 124, p. 13680-13681.

Nishio, M.: The CH/[small pi] hydrogen bond in chemistry. Conformation, supramolecules, optical resolution and interactions involving carbohydrates. *Physical Chemistry Chemical Physics* **2011**, 13, 13873-13900.

Nishio, M. et al. CH/ $\pi$  hydrogen bonds in organic and organometallic chemistry. *CrystEngComm* **2012**, v. 11, n. 9, p. 1757-1788.

Nunn, C. M.; Jenkins, T. C.; Neidle, S. Crystal structure of d(CGCGAATTCGCG) complexed with propamidine, a short-chain homologue of the drug pentamidine. *Biochemistry* **1993**, v. 32, p. 13838-13843.

\_\_\_\_\_. Crystal structure of gamma-oxapentamidine complexed with d(CGCGAATTCGCG)<sub>2</sub>. The effects of drug structural change on DNA minor-groove recognition. *European Journal of Biochemistry* **1994**, v. 226, p. 953-961.

Pauling, L.; Corey, R. B.; Branson, H. R. The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chain. *Proceedings of the National Academy of Sciences* 1951, v. 37, n. 4, p. 205-211.

Pinheiro, S. et al. Structural and energetic study of cation-[small pi]-cation interactions in proteins. *Physical Chemistry Chemical Physics* **2017**, v. 19, n. 15, p. 9849-9861.

Pjura, P. E.; Grzeskowiak, K.; Dickerson, R. E. Binding of Hoechst 33258 to the minor groove of B-DNA. *Journal of Molecular Biology* **1987**, v. 197, p. 257-271.

Protein Data Bank (PDB): [www.rcsb.org/pdb](http://www.rcsb.org/pdb) , acesso em 28-03-2010.

Protein Data Bank Summary (PDBsum): <http://www.ebi.ac.uk/thornton-srv/databases/pdbsum/> , acesso em 28-03-2010.

Pullman, B. Specificity in the interaction of non intercalative groove binding ligands with nucleic acids. Proc. Int. Symp. Biomol. Struct. Interactions, Suppl. J. Biosci. India **1985**. 8: p. 681-688.

Quigley, G. J. et al. Non-Watson-Crick G-C and A-T base pairs in a DNA-antibiotic complex. *Science* **1986**, v. 232, p. 1255-1259.

Quintana, J. R.; Lipanov, A. A.; Dickerson, R. E. Low-temperature crystallographic analyses of the binding of Hoechst 33258 to the double-helical DNA dodecamer C-G-C-G-A-A-T-T-C-G-C-G. *Biochemistry* **1991**, v. 30, p. 10294-10306.

RCSB. RCSB Protein Data Bank - RCSB PDB. 09/09/2012 2012. [http://pdbeta.rcsb.org/pdb/static.do?p=general\\_information/pdb\\_statistics/index.html](http://pdbeta.rcsb.org/pdb/static.do?p=general_information/pdb_statistics/index.html) , acesso em 30-11-2012.

Sacco, A. C. S.: Mapeamento vetorial de interações eletrostáticas para biomoléculas. **2013**. 156 (Mestrado). Programa de Pós-Graduação em Biotecnologia, UFSCar, São Carlos.

Sacco, A. C. S.; Caracelli, I.; Zukerman-Schpector, J.; Souza, L. M.: WIM - Weak Interacion Mapping. Tatuí **2016**.

Sacco, A. C. S.; Zuckerman-Schpector, J.; Caracelli, I. Aplicação do mapeamento vetorial em DNA. V Simpósio de Biotecnologia da Universidade Federal de São Carlos, **2016**, São Carlos. PPG Biotec - UFSCar. p.5-6.

Sacco, A. C. S.; Caracelli, I.; Zukerman-Schpector, J. Mapeamento vetorial de interações eletrostáticas para macromoléculas., 1.º Encontro Nacional da Rede Nacional de Nanobiotecnologia Aplicada à Medicina e à Defesa, **2010**, São Carlos. PPG Biotec - UFSCar. p.7.

Schweitzer, B. I. et al. Solution structure of a DNA dodecamer containing the anti-neoplastic agent arabinosylcytosine: combined use of NMR, restrained molecular dynamics, and full relaxation matrix refinement. *Biochemistry* **1994**, v. 33, p. 11460-11475.

Scolnick, L. R. et al. Novel Binding Mode of Hydroxamate Inhibitors to Human Carbonic Anhydrase II. *Journal of the American Chemical Society* **1997**, v. 119, n. 4, p. 850-851.

Sebesta, R. W. Concepts of programming languages. 10. Boston: Addison-Wesley, 2012. 795.

Seng, H.-L. et al. Factors affecting nucleolytic efficiency of some ternary metal complexes with DNA binding and recognition domains. Crystal and molecular structure of Zn(phen)(edda). *Journal of Inorganic Biochemistry* **2008**, v. 102, n. 11, p. 1997-2011.

Seng, H.-L. et al. Crystal structure, DNA binding studies, nucleolytic property and topoisomerase I inhibition of zinc complex with 1,10-phenanthroline and 3-methyl-picolinic acid. *BioMetals* **2010**, v. 23, n. 1, p. 99-118.

Seng, H.-L.; Von, S.-T.; Tan, K.-W.; Maah, M.; Ng, S.-W.; R., Raja; Caracelli, I.; Ng, C.-H.: Crystal structure, DNA binding studies, nucleolytic property and topoisomerase I inhibition of zinc complex with 1,10-phenanthroline and 3-methyl-picolinic acid. *BioMetals* **2010**, 23, 99-118.

Shaikh, S. A.; Jayaram, B. DNA Drug Interaction. New Delhi, India: Supercomputing Facility for Bioinformatics & Computational Biology, IIT Delhi 2012.

Shanthi, V.; Ramanathan, K.; Sethumadhavan, R.: Exploring the Role of C-H $\cdots$  $\pi$  Interactions on the Structural Stability of Single Chain "All-Alpha" Proteins. *Applied Biochemistry and Biotechnology* **2010**, 160, 1473-1483.

Sharp, J. Visual C# 2010 step by step. Redmond: Microsoft Press, 2010.

Shatzky-Schwartz, M. et al. X-ray and solution studies of DNA oligomers and implications for the structural basis of A-tract-dependent curvature. *Journal of Molecular Biology* **1997**, v. 267, p. 595-623.

Sherrill, C. D.; Sumpter, B. G.; Sinnokrot, M. O.; Marshall, M. S.; Hohenstein, E. G.; Walker, R. C.; Gould, I. R.: Assessment of standard force field models against high-quality ab initio potential curves for prototypes of  $\pi$ - $\pi$ , CH/ $\pi$ , and SH/ $\pi$  interactions. *Journal of Computational Chemistry* **2009**, 30, 2187-2193.

Shui, X. et al. The B-DNA dodecamer at high resolution reveals a spine of water on sodium. *Biochemistry* **1998**, v. 37, p. 8341-8355.

\_\_\_\_\_. Structure of the potassium form of CGCGAATTCGCG: DNA deformation by electrostatic collapse around inorganic cations. *Biochemistry* **1998**, v. 37, p. 16877-16887.

Simpson, I. J. et al. DNA minor groove interactions and the biological activity of 2, 5-bis-[4-(N-alkylamidino) phenyl] furans. *Bioorganic & Medicinal Chemistry Letters* **2000**, v. 10, n. 23, p. 2593-2597.

Sines, C. C. et al. Cations Mediate B-DNA Conformational Heterogeneity. *Journal of the American Chemical Society* **2000**, v. 122, p. 11048-11056.

Singh, N. J.; Min, S. K.; Kim, D. Y.; Kim, K. S.: Comprehensive Energy Analysis for Various Types of  $\pi$ -Interaction. *Journal of Chemical Theory and Computation* **2009**, 5, 515-529.

Smith, T.; Slipchenko, L. V.; Gordon, M. S. Modeling  $\pi$ - $\pi$  Interactions with the Effective Fragment Potential Method: The Benzene Dimer and Substituents. *The Journal of Physical Chemistry A* **2008**, v. 112, n. 23, p. 5286-5294.

Squire, C. J. et al. Structures of m-iodo Hoechst-DNA complexes in crystals with reduced solvent content: implications for minor groove binder drug design. *Nucleic Acids Research* **2000**, v. 28, n. 5, p. 1252-1258.

Squire, C. J.; Clark, G. R.; Denny, W. A. Minor groove binding of a bis-quaternary ammonium compound: the crystal structure of SN 7167 bound to d(CGCGAATTCGCG)<sub>2</sub>. *Nucleic Acids Research* **1997**, v. 25, p. 4072-4078.

Sriram, M. et al. Conformation of B-DNA containing O6-ethyl-G-C base pairs stabilized by minor groove binding drugs: molecular structure of d(CGC[e6G]AATTCGCG) complexed with Hoechst 33258 or Hoechst 33342. *The EMBO Journal* **1992a**, v. 11, p. 225-232.

\_\_\_\_\_. Structural consequences of a carcinogenic alkylation lesion on DNA: effect of O6-ethylguanine on the molecular structure of the d(CGC[e6G]AATTCGCG)-netropsin complex. *Biochemistry* **1992b**, v. 31, p. 11823-11834.

Steinbruch, A.; Winterle, P.: Álgebra linear e geometria analítica. São Paulo: Pearson Prentice Hall, **1987**. 470 p.

Steiner, T. The Hydrogen Bond in the Solid State. *Angewandte Chemie International Edition* **2002**, v. 41, n. 1, p. 48-76.

Steiner, T.; Koellner, G. Hydrogen bonds with  $\pi$ -acceptors in proteins: frequencies and role in stabilizing local 3D structures. *Journal of Molecular Biology* **2001**, v. 305, n. 3, p. 535-557.

Takahashi, O.; Kohno, Y.; Nishio, M. Relevance of weak hydrogen bonds in the conformation of organic compounds and bioconjugates: Evidence from recent experimental data and high-level ab initio mo calculations. *Chemical Reviews* **2010**, v. 110, n. 10, p. 6049-6076.

Teng, M. K. et al. The molecular structure of the complex of Hoechst 33258 and the DNA dodecamer d(CGCGAATTCGCG). *Nucleic Acids Research* **1988**, v. 16, p. 2671-2690.

Tereshko, V.; Minasov, G.; Egli, M. A. *Journal of the American Chemical Society* **1999**, v. 121, p. 3590-3595.

Tewari, A. K.; Dubey, R. Emerging trends in molecular recognition: Utility of weak aromatic interactions. *Bioorganic & Medicinal Chemistry* **2008**, 16, 126-143.

Tiekink, E. R. T.; Zukerman-Schpector, J. A structural survey of metal [small pi] heteroaromatic supramolecular synthons for metal = tellurium, tin, and gold. *CrystEngComm* 2009, v. 11, n. 12, p. 2701-2711.

Tiekink, E. R. T.; Zukerman-Schpector, J.: Gold[small pi] aryl interactions as supramolecular synthons. *CrystEngComm* **2009**, 11, 1176-1186.

\_\_\_\_\_. Emerging supramolecular synthons: C-H [small pi](chelate) interactions in metal bis(1,1-dithiolates). *Chemical Communications* **2001**, v. 47, n. 23, p. 6623-6625.

Tjandra, N. et al. The NMR Structure of a DNA Dodecamer in an Aqueous Dilute Liquid Crystalline Phase. *Journal of the American Chemical Society* **2000**, v. 122, p. 6190-6200.

Tomić, Z. D.; Leovac, V. M.; Jaćimović, Ž K.; Giester, G.; Zarić, S. D.: Multi- $\pi$ ,  $\pi$  stacked interaction between planar chelate rings in the crystal structure of dichloro(2-hydroxy-1-naphthaldehyde-3-methylisothiosemicarbazonato)iron(III) hemiethanol solvate. *Inorganic Chemistry Communications* **2006**, 9, 833-835.

Tsuzuki, S.; Honda, K.; Uchamaru, T.; Mikami, M.; Tanabe, K.: Origin of the Attraction and Directionality of the NH/ $\pi$  Interaction: Comparison with OH/ $\pi$  and CH/ $\pi$  Interactions. *Journal of the American Chemical Society* **2000**, 122, 11450-11458.

Vega-Teijido, M.; Caracelli, I.; Zukerman-Schpector, J.: Conformational analyses and docking studies of a series of 5-nitrofuranyl- and 5-nitrothiophenyl-semicarbazone derivatives in three possible binding sites of trypanothione and glutathione reductases. *Journal of Molecular Graphics and Modelling* **2006**, 24, 349-355.

Watson, J. D.; Crick, F. H. Molecular structure of nucleic acids. *Nature* 1953, v. 171, n. 4356, p. 737-738.

Westhof, E. Re-refinement of the B-dodecamer d(CGCGAATTCGCG) with a comparative analysis of the solvent in it and in the Z-hexamer d(5BrCG5BrCG5BrCG). *Journal of Biomolecular Structure and Dynamics* **1987**, v. 5, p. 581-600.

Wellenzohn, B. et al. Significance of Ligand Tails for Interaction with the Minor Groove of B-DNA. *Biophysical Journal* **2001**, v. 81, n. 3, p. 1588-1599.

Wheeler, S. E.; McNeil, A. J.; Müller, p.; Swager, T. M.; Houk, K. N.: Probing Substituent Effects in Aryl–Aryl Interactions Using Stereoselective Diels–Alder Cycloadditions. *Journal of the American Chemical Society* **2010**, 132, 3304-3311.

Williams, L. D.; Maher III, L. J. Electrostatic Mechanisms of DNA deformation. *Annual Review of Biophysics and Biomolecular Structure* **2000**, v. 29, n. 1, p. 497-521.

WING, R. M. et al. The primary mode of binding of cisplatin to a B-DNA dodecamer: C-G-C-G-A-A-T-T-C-G-C-G. *The EMBO Journal* **1984**, v. 3, p. 1201-1206.

Wintjens, R.; Liévin, J.; Rooman, M.; Buisine, E.: Contribution of cation- $\pi$  interactions to the stability of protein-DNA complexes. *Journal of Molecular Biology* **2000**, 302, 393-408.

Wood, A. A. et al. Variability in DNA minor groove width recognised by ligand binding: the crystal structure of a bis-benzimidazole compound bound to the DNA duplex d(CGCGAATTCGCG)<sub>2</sub>. *Nucleic Acids Research* **2000**, v. 23, p. 3678-3684.

Woods, K. K. et al. Monovalent Cations Sequester within the A-Tract Minor Groove of [d(CGCGAATTCGCG)]<sub>2</sub>. *Journal of the American Chemical Society* **2000**, v. 122, p. 1546-1547.

Wu, Z. et al. Overall structure and sugar dynamics of a DNA dodecamer from homo- and heteronuclear dipolar couplings and (31)P chemical shift anisotropy. *Journal of Biomolecular NMR* **2003**, v. 26, p. 297-315.

Zhang, Z.; Wu, Y.; Zhang, G.: N[small pi] and O[small pi] interactions rather than "directed" hydrogen bonding in the crystal packing of ethyl (Z)-2-cyano-3-[(4-alkoxyphenyl)amino]prop-2-enoates. *CrystEngComm* **2011**, 13, 4496-4499.

Zhu, W. et al. Antibacterial drug leads: DNA and enzyme multitargeting. *Journal of medicinal chemistry* **2015**, v. 58, n. 3, p. 1215-1227.

Zukerman-Schpector, J.; Caracelli, I.; Haiduc, I. LaCrEMM - Laboratório de Cristalografia Estereodinâmica e Modelagem Molecular. **2012**.

Zukerman-Schpector, J.; Haiduc, I.: Tellurium[small pi]-aryl interactions: a new bonding motif for supramolecular self-assembly and crystal engineering. *CrystEngComm* **2002**, 4, 178-193.