

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

**“MAPEAMENTO VETORIAL DE  
INTERAÇÕES ELETROSTÁTICAS PARA  
BIOMOLÉCULAS”**

**ANTONIO CÉSAR SILVA SACCO**

**São Carlos  
2013**

ANTONIO CÉSAR SILVA SACCO\*

**“MAPEAMENTO VETORIAL DE  
INTERAÇÕES ELETROSTÁTICAS PARA BIOMOLÉCULAS”**

Monografia apresentada ao  
Programa de Mestrado em  
Biotecnologia do Programa de Pós  
Graduação em Biotecnologia da  
Universidade Federal de São Carlos.

Orientadores: **Profa. Dra. Ignez Caracelli**  
**Prof. Dr. Julio Zukerman-Schpector**

**\* Bolsista CAPES**

São Carlos  
2013

**Ficha catalográfica elaborada pelo DePT da  
Biblioteca Comunitária da UFSCar**

S119mv

Sacco, Antonio César Silva.

Mapeamento vetorial de interações eletrostáticas para biomoléculas / Antonio César Silva Sacco. -- São Carlos : UFSCar, 2013.

159 f.

Acompanha anexo em CD-ROM.

Dissertação (Mestrado) -- Universidade Federal de São Carlos, 2013.

1. Biotecnologia. 2. Mapeamento vetorial. 3. Interações não covalentes. 4. Interações eletrostáticas. 5. Interações pi.  
I. Título.

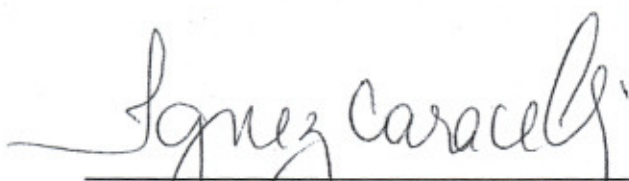
CDD: 660.6 (20<sup>a</sup>)

Antonio César Silva Sacco

Dissertação de Mestrado submetida  
à Coordenação do Programa de  
Pós-Graduação em Biotecnologia,  
da Universidade Federal de São  
Carlos, como requisito parcial para  
a obtenção do título de Mestre em  
Biotecnologia

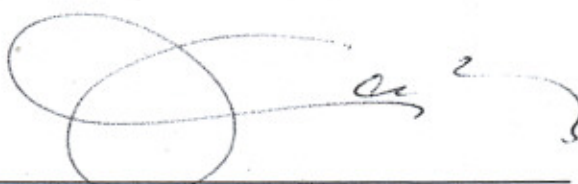
Aprovado em: 15/04/2013

**BANCA EXAMINADORA**



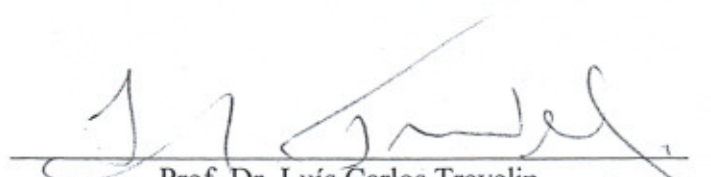
---

Prof.<sup>a</sup> Dr.<sup>a</sup> Ignez Caracelli (Orientadora)  
(Universidade Federal de São Carlos)



---

Prof. Dr. Carlos A. de Simone  
(Universidade Federal de Alagoas)



---

Prof. Dr. Luís Carlos Trevelin  
(Universidade Federal de São Carlos)

*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 a minha presença ao convívio familiar.*

*Aos meus pais, Clemente Sacco e Lelia 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 cuja humildade compartilha a sapiência e é uma verdadeira honra ser integrante de sua equipe.*

*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, não esquecendo os agradáveis momentos de trabalho, de descontração e companheirismo.*

*Ao meu amigo, Prof. Dr. Fernando Manuel Araújo Moreira pela acolhida ao Programa de Pós Graduação em Biotecnologia da UFSCar.*

*À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela Bolsa concedida.*

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

Sacco, Antonio César Silva. **Mapeamento vetorial de interações eletrostáticas para biomoléculas**. 2013. Dissertação (Programa de Pós-Graduação em Biotecnologia) – UFSCar, São Carlos, 2013.

## RESUMO

As interações não covalentes envolvendo anéis aromáticos, sistemas  $\pi$ , possuem um papel importante em uma ampla gama de complexos. A sua compreensão é essencial para o desenho de drogas já que a maioria das estruturas cristalinas de complexos de proteínas com moléculas pequenas revela a existência destas interações envolvendo as cadeias laterais de aminoácidos aromáticos no receptor e/ou anéis aromáticos e hetero-aromáticos dos ligantes atuando como receptores ou doadores.

A análise dos parâmetros estruturais envolvidos nas interações  $\pi$  começa com a determinação e mapeamento dos centros geométricos (centroides) dos grupos aromáticos e segue com o cálculo de distâncias, ângulos e ângulos diedros a átomos e/ou outros sistemas  $\pi$  dentro de determinados limites pré-estabelecidos. Isto é combinado com uma análise em tela gráfica dos resultados. Desta forma é possível determinar as variáveis necessárias à criação de um modelo matemático para o mapeamento das interações  $\pi$  e criar uma ferramenta para localizá-las e assim auxiliar na identificação de possíveis sítios de ligação a fármacos ou outros ligantes, bem como criar “graph sets” para comparar duas macromoléculas distintas, porém funcionalmente semelhantes.

Conseqüentemente será possível ter um software que automatizará o trabalho de identificação das interações  $\pi$ , reduzindo o tempo da pesquisa e proporcionando ao pesquisador um aumento da sua produtividade.

**Palavras-chave:** mapeamento vetorial, interações  $\pi$ , interações eletrostáticas, prospecção e análise de interações  $\pi$ .

Sacco, Antonio César Silva. **Vector mapping of electrostatic interactions to biomolecules.** 2013. Dissertation (Postgraduate Program in Biotechnology) – UFSCar, São Carlos, 2013.

## **ABSTRACT**

Non-covalent interactions involving aromatic rings  $\pi$ , systems play an important role in a wide range of complexes. Their understanding is essential for drug design since most of the crystal structures of complexes of proteins with small molecules reveal the existence of these interactions involving the side chains of aromatic amino acids in the receptor and / or aromatic rings and hetero-aromatic ligands acting as acceptors or donors.

The analysis of the structural parameters involved in  $\pi$  interactions begins with the determination and mapping of the geometrical centers (centroid) of the aromatic groups followed by the calculation of distances, angles and dihedral angles to atoms and / or other  $\pi$  systems, within pre-established limits. This is combined with a visual analysis of the results in graphical displays. In this way it is possible to determine the variables needed to create a mathematical model for mapping the possible  $\pi$ -interactions and to create a tool to locate them. This will assist in identifying possible binding sites of the drugs or other ligands and to create "graph sets" to compare two functionally similar but different macromolecules.

As a result, we have a software that will automate the job of identifying  $\pi$ -interactions, reducing search time and providing the researcher to increase their productivity.

**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 gráfica de parte do enovelamento de uma glutatona redutase (1GRA), onde podem ser observadas as interações $O \cdots \pi$ entre a PHE248 e a HIS219 e a $N \cdots \pi$ entre a HIS219 e a ARG224.	23
<b>Figura 1.2</b>	Representação gráfica de parte de uma molécula de DNA com as interações $\pi$ existentes entre as bases.	23
<b>Figura 1.3</b>	Representação esquemática dos parâmetros geométricos envolvidos em uma interação $\pi$ .	24
<b>Figura 1.4</b>	Representação da distribuição de cargas do hexafluoreno e benzeno.	26
<b>Figura 1.5</b>	Representação esquemática de uma tirosina e sua interação $\pi$ , com o dipolo da nuvem eletrônica sendo induzido em função de uma carga.	28
<b>Figura 1.6</b>	Representações gráficas dos tipos de interação $\pi \cdots \pi$ : <b>(a)</b> face a face, <b>(b)</b> paralela-deslocada, <b>(c)</b> forma de "T".	29
<b>Figura 1.7</b>	Interações $\pi \cdots \pi$ .	30
<b>Figura 1.8</b>	Interações $\pi \cdots \pi$ esquema de distribuição de cargas.	31
<b>Figura 1.9</b>	Parâmetros geométricos para as interações $\pi \cdots \pi$ .	32
<b>Figura 1.10</b>	Representação gráfica de uma interação $\pi \cdots \pi$ entre as bases DA 18 e DT 19 de uma molécula de DNA.	33
<b>Figura 1.11</b>	Representação esquemática de uma interação $\pi \cdots \sigma$ , considerando como doador o átomo de um centroide ligado ao hidrogênio e acceptor a nuvem eletrônica do segundo anel aromático.	34
<b>Figura 1.12</b>	Representação gráfica de uma interação C-H $\cdots \pi$ T-shaped.	35
<b>Figura 1.13</b>	Representação gráfica de uma interação N-H $\cdots \pi$ entre duas bases de uma molécula de DNA.	36
<b>Figura 1.14</b>	Representação gráfica de uma interação O-H $\cdots \pi$ em uma cyclooxygenase-2.	36
<b>Figura 1.15</b>	Nuvem eletrônica de um anel aromático que atua como doadora de elétrons para uma interação $\text{cátion} \cdots \pi$ .	38
<b>Figura 1.16</b>	Esquema de uma interação $\text{cátion} \cdots \pi$ , representando-se um cátion genérico posicionado sobre a face de um anel de benzeno.	38
<b>Figura 1.17</b>	Representação gráfica de uma interação $N \cdots \pi$ entre as bases G 12 e G 14 de uma molécula de DNA.	39

<b>Figura 1.18</b>	Representação gráfica de uma interação $O \cdots \pi$ entre a HIS219 e a PHE248.	39
<b>Figura 1.19</b>	Representação esquemática do padrão de difração de um cristal de lisozima.	42
<b>Figura 1.20</b>	Representação do processo de cristalografia por difração de raios-X.	42
<b>Figura 1.21</b>	Partes de registros de um arquivo de dados PDB.	44
<b>Figura 1.22</b>	Estatística referente às estruturas depositadas no PDB até a data de 04/09/2012.	45
<b>Figura 1.23</b>	Representação esquemática das estruturas primárias, secundárias e terciárias de uma proteína.	46
<b>Figura 1.24</b>	Representação da série de Fourier para a mioglobina tipo A.	46
<b>Figura 1.25</b>	Fotografia com vista parcial do EDSAC.	47
<b>Figura 1.26</b>	Ângulos $\phi$ (phi) e $\psi$ (psi) entre ligações peptídicas.	48
<b>Figura 1.27</b>	Estrutura de DNA sendo modelada pelo computador.	49
<b>Figura 1.28</b>	Formato de banco de dados de parte de uma proteína obtido por cristalografia	50
<b>Figura 1.29</b>	Representação esquemática do volume de controle criado pelo cubo de arestas $a$ , $b$ e $c$ .	51
<b>Figura 1.30</b>	Output do programa do computador pela análise seqüencial dos átomos de uma molécula de mioglobina.	52
<b>Figura 1.31</b>	Representação de um banco de dados unidimensional.	53
<b>Figura 1.32</b>	Estrutura de banco de dados para moléculas.	55
<b>Figura 1.33</b>	Banco de dados com os átomos padrão para a construção de moléculas e suas respectivas classificações de ligações.	56
<b>Figura 1.34</b>	Banco de dados com os ângulos das ligações para os átomos.	56
<b>Figura 1.35</b>	Banco de dados de grupos pré-definidos para a modelagem molecular.	57
<b>Figura 1.36</b>	Representação química da sequência $CH_3.CH_2.CO.CH.C.H_2^*$ .	57
<b>Figura 1.37</b>	Representação de uma molécula no formato de árvore binária.	58
<b>Figura 1.38</b>	Imagem capturada de uma tela com o programa RasMol.	60
<b>Figura 1.39</b>	Imagens da lisozima 1HEL capturada de uma tela com o programa Chime.	61
<b>Figura 1.40</b>	Visualização on-line da estrutura tridimensional da proteína glutathione reductase 1GRE utilizando o Jmol.	62
<b>Figura 1.41</b>	Visualização da estrutura tridimensional de um DNA com as interações $\pi \cdots \pi$ , utilizando o software Discovery Studio 3.1.	63
<b>Figura 1.42</b>	Caixa de diálogo para alteração dos atributos envolvidos no cálculo das interações $\pi \cdots \pi$ .	64
<b>Figura 1.43</b>	Empilhamento de dois anéis envolvidos em uma interação $\pi \cdots \pi$ .	65

<b>Capítulo 2 – Materiais e métodos</b>		<b>página</b>
<b>Figura 2.1</b>	Representação gráfica de uma glutathiona redutase 1GRA com seus átomos.	69
<b>Figura 2.2</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.	70
<b>Figura 2.3</b>	Sequencia de poses de software de visualização gráfica onde pode ser visto o trabalho de busca de uma possível interação $\pi$ entre o centroide do resíduo de aminoácido TYR 23 do monômero A de uma glutathiona redutase 1GRA com todos os átomos adjacentes.	71
<b>Figura 2.4</b>	Localização dos átomos envolvidos no cálculo dos centroides de um triptofano.	74
<b>Figura 2.5</b>	Determinação do vetor normal ao plano do anel aromático de uma tirosina.	75
<b>Figura 2.6</b>	Determinação do ângulo $\alpha$ entre o vetor normal ao plano $\vec{n}$ e o vetor distância $\vec{v}$ com base na estrutura de uma tirosina.	77
<b>Figura 2.7</b>	Fluxograma das operações envolvidas no mapeamento vetorial para as interações $\pi$ .	79
<b>Figura 2.8</b>	Imagem da planilha eletrônica construída para o cálculo da coordenada do centroide da base Citosina.	80
<b>Figura 2.9</b>	Planilha com a relação das coordenadas dos centroides e vetores normais ao plano de cada anel aromático, calculados em um DNA sem ligante.	81
<b>Figura 2.10</b>	Visualização dos centroides calculados em um DNA sem ligante.	82
<b>Figura 2.11</b>	Visualização dos centroides calculados sobrepostos à estrutura determinada de um DNA sem ligante.	83
<b>Figura 2.12</b>	Imagem da planilha eletrônica construída 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.	84
<b>Figura 2.13</b>	Imagem da planilha eletrônica com os resultados válidos do módulo do vetor distância, onde são excluídos os resultados envolvendo átomos e centroides da mesma base, bem como átomos de carbono.	85
<b>Figura 2.14</b>	Imagem da planilha eletrônica com os resultados válidos para o ângulo relativo ao vetor normal.	86
<b>Figura 2.15</b>	Imagem da planilha eletrônica utilizada para exportar os resultados válidos em uma interação $\pi$ , na em arquivo compatível para visualização em software gráfico.	87
<b>Figura 2.16</b>	Visualização da imagem da interação $N^{\cdots}\pi$ obtida por meio de cálculo vetorial em um DNA sem ligantes.	87
<b>Figura 2.17</b>	Visualização da interação $N^{\cdots}\pi$ obtida por meio de cálculo vetorial sobre as bases de um DNA sem ligantes.	88

<b>Capítulo 3 – Resultados e Discussões</b>		<b>página</b>
<b>Figura 3.1</b>	Representação gráfica das interações $\pi$ mapeadas vetorialmente para a proteína código PDB 1GRA	91
<b>Figura 3.2</b>	Imagens das estruturas de DNA com as interações $\pi$ identificadas por mapeamento vetorial.	94
<b>Figura 3.3</b>	Imagens das estruturas de DNA com as interações $\pi$ identificadas por mapeamento vetorial, excluídas as interações entre o ligante e a base.	96
<b>Figura 3.4</b>	Representação gráfica das interações $\pi$ entre as bases e o ligante de uma estrutura de DNA com o ligante intercalado.	113
<b>Figura 3.5</b>	Representação gráfica das interações $\pi$ entre o ligante e as bases de uma estrutura de DNA com o ligante posicionado no sulco.	114
<b>Figura 3.6</b>	Visualização gráfica da deformação calculada em função da presença de ligante no sulco de estrutura de DNA.	116
<b>Figura 3.7</b>	Visualização gráfica do deslocamento dos centroides em função da presença de ligante no sulco de estrutura de DNA.	116
<b>Figura 3.8</b>	Visualização gráfica da deformação calculada em função da presença de ligante intercalado na estrutura de DNA.	118
<b>Figura 3.9</b>	Visualização gráfica do deslocamento dos centroides em função da presença do ligante intercalado à estrutura de DNA.	118
<b>Figura 3.10</b>	Diagrama de transferência de cargas eletrostáticas nos sistemas $\pi$ nas estruturas de DNA.	119
<b>Figura 3.11</b>	Diagrama de transferência de cargas eletrostáticas nas interações $N^{\cdots}\pi$ nas estruturas de DNA.	120
<b>Figura 3.12</b>	Sobreposição das interações $N^{\cdots}\pi$ das estruturas de DNA com ligantes sobre a estrutura sem ligante.	121
<b>Figura 3.13</b>	Diagrama de transferência de cargas eletrostáticas nas interações $O^{\cdots}\pi$ nas estruturas de DNA.	122
<b>Figura 3.14</b>	Sobreposição das interações $O^{\cdots}\pi$ das estruturas de DNA com ligantes sobre a estrutura sem ligante	123
<b>Figura 3.15</b>	Diagrama de transferência de cargas eletrostáticas nas interações $\pi^{\cdots}\pi$ nas estruturas de DNA.	124
<b>Figura 3.16</b>	Sobreposição das interações $\pi^{\cdots}\pi$ das estruturas de DNA com ligantes sobre a estrutura sem ligante	125
<b>Figura 3.17</b>	Diagrama de transferência de cargas eletrostáticas nas interações $C-H^{\cdots}\pi$ nas	126

	estruturas de DNA.	
<b>Figura 3.18</b>	Sobreposição das interações C-H $\cdots$ $\pi$ das estruturas de DNA com ligantes sobre a estrutura sem ligante.	127
<b>Figura 3.19</b>	Diagrama de transferência de cargas eletrostáticas nas interações N-H $\cdots$ $\pi$ nas estruturas de DNA.	128
<b>Figura 3.20</b>	Sobreposição das interações N-H $\cdots$ $\pi$ das estruturas de DNA com ligantes sobre a estrutura sem ligante.	129
<b>Figura 3.21</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.	130
<b>Figura 3.22</b>	Representação gráfica da tabulação das distâncias das interações N $\cdots$ $\pi$ das estruturas de DNA com ligantes.	131
<b>Figura 3.23</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 de DNA com ligantes.	131
<b>Figura 3.24</b>	Representação gráfica da tabulação das distâncias das interações O $\cdots$ $\pi$ das estruturas de DNA com ligantes.	132
<b>Figura 3.25</b>	Representação gráfica da tabulação dos ângulos entre o vetor distância e o vetor normal ao plano do anel aromático das interações O $\cdots$ $\pi$ das estruturas de DNA com ligantes.	132
<b>Figura 3.26</b>	Representação gráfica da tabulação das distâncias das interações $\pi\cdots\pi$ das estruturas de DNA com ligantes.	133
<b>Figura 3.27</b>	Representação gráfica da tabulação dos ângulos entre o vetor distância e o vetor normal ao plano do anel aromático das interações $\pi\cdots\pi$ das estruturas de DNA com ligantes.	133
<b>Figura 3.28</b>	Representação gráfica da tabulação das distâncias das interações C-H $\cdots$ $\pi$ das estruturas de DNA com ligantes.	134
<b>Figura 3.29</b>	Representação gráfica da tabulação dos ângulos entre o vetor distância e o vetor normal ao plano do anel aromático das interações C-H $\cdots$ $\pi$ das estruturas de DNA com ligantes.	134
<b>Figura 3.30</b>	Representação gráfica da tabulação das distâncias das interações N-H $\cdots$ $\pi$ das estruturas de DNA com ligantes.	135
<b>Figura 3.31</b>	Representação gráfica da tabulação dos ângulos entre o vetor distância e o vetor normal ao plano do anel aromático das interações N-H $\cdots$ $\pi$ das estruturas de DNA com ligantes.	135

## LISTA DE TABELAS

<b>Capítulo 1 - Introdução</b>		<b>página</b>
<b>Tabela 1.1</b>	Forças eletrostáticas simples.	28
<b>Tabela 1.2</b>	Forças eletrostáticas combinadas.	29
<b>Tabela 1.3</b>	Distribuição das estruturas de macromoléculas por tipo depositadas no PDB.	41
<b>Tabela 1.4</b>	Métodos utilizados para a obtenção das estruturas de macromoléculas depositadas no PDB.	41
<b>Tabela 1.5</b>	Formatos de arquivos processados pelo Jmol.	62
 <b>Capítulo 2 – Materiais e métodos</b>		 <b>página</b>
<b>Tabela 2.1</b>	Fragmento de estrutura cristalográfica de uma glutatona redutase, especificamente a Tyr 23.	72
<b>Tabela 2.2</b>	Parâmetros geométricos para as interações $\pi$ , por tipo.	73
 <b>Capítulo 3 – Resultados e Discussões</b>		 <b>página</b>
<b>Tabela 3.1</b>	Comparativo de interações $\pi$ por tipo: proteína e DNA.	90
<b>Tabela 3.2</b>	Comparativo das interações $\pi$ identificadas por estrutura de DNA, incluídas as interações presentes nos ligantes.	93
<b>Tabela 3.3</b>	Comparativo percentual das interações $\pi$ identificadas por estrutura de DNA, incluídas as interações presentes nos ligantes.	94
<b>Tabela 3.4</b>	Comparativo das interações $\pi$ identificadas por estrutura de DNA, não computadas as interações presentes nos ligantes.	95
<b>Tabela 3.5</b>	Comparativo percentual das interações $\pi$ identificadas por estrutura de DNA, não computadas as interações presentes nos ligantes.	96
<b>Tabela 3.6</b>	Demonstrativo das interações $\pi$ identificadas por estrutura de DNA existentes entre as bases e ligantes.	97
<b>Tabela 3.7</b>	Demonstrativo das interações $\pi$ com origem na base C1 com os átomos e centroides relacionados por bases.	97
<b>Tabela 3.8</b>	Demonstrativo das interações $\pi$ com origem na base G2, anel aromático com 6 átomos, com os átomos e centroides relacionados por bases.	98
<b>Tabela 3.9</b>	Demonstrativo das interações $\pi$ com origem na base G2, anel aromático com 5 átomos, com os átomos e centroides relacionados por bases.	99

<b>Tabela 3.10</b>	Demonstrativo das interações $\pi$ com origem na base C3, com os átomos e centroides relacionados por bases.	99
<b>Tabela 3.11</b>	Demonstrativo das interações $\pi$ com origem na base G4, anel aromático com 6 átomos, com os átomos e centroides relacionados por bases.	100
<b>Tabela 3.12</b>	Demonstrativo das interações $\pi$ com origem na base G4, anel aromático com 5 átomos, com os átomos e centroides relacionados por bases.	100
<b>Tabela 3.13</b>	Demonstrativo das interações $\pi$ com origem na base A5, anel aromático com 6 átomos, com os átomos e centroides relacionados por bases.	100
<b>Tabela 3.14</b>	Demonstrativo das interações $\pi$ com origem na base A5, anel aromático com 5 átomos, com os átomos e centroides relacionados por bases.	101
<b>Tabela 3.15</b>	Demonstrativo das interações $\pi$ com origem na base A6, anel aromático com 6 átomos, com os átomos e centroides relacionados por bases.	101
<b>Tabela 3.16</b>	Demonstrativo das interações $\pi$ com origem na base A6, anel aromático com 5 átomos, com os átomos e centroides relacionados por bases.	102
<b>Tabela 3.17</b>	Demonstrativo das interações $\pi$ com origem na base T7, com os átomos e centroides relacionados por bases.	102
<b>Tabela 3.18</b>	Demonstrativo das interações $\pi$ com origem na base T8, com os átomos e centroides relacionados por bases.	103
<b>Tabela 3.19</b>	Demonstrativo das interações $\pi$ com origem na base C9, com os átomos e centroides relacionados por bases.	103
<b>Tabela 3.20</b>	Demonstrativo das interações $\pi$ com origem na base G10, anel aromático com 6 átomos, com os átomos e centroides relacionados por bases.	103
<b>Tabela 3.21</b>	Demonstrativo das interações $\pi$ com origem na base G10, anel aromático com 5 átomos, com os átomos e centroides relacionados por bases.	104
<b>Tabela 3.22</b>	Demonstrativo das interações $\pi$ com origem na base C11, com os átomos e centroides relacionados por bases.	104
<b>Tabela 3.23</b>	Demonstrativo das interações $\pi$ com origem na base G12, anel aromático com 6 átomos, com os átomos e centroides relacionados por bases.	104
<b>Tabela 3.24</b>	Demonstrativo das interações $\pi$ com origem na base G12, anel aromático com 5 átomos, com os átomos e centroides relacionados por bases.	105
<b>Tabela 3.25</b>	Demonstrativo das interações $\pi$ com origem na base C13, com os átomos e centroides relacionados por bases.	105
<b>Tabela 3.26</b>	Demonstrativo das interações $\pi$ com origem na base G14, anel aromático com	105

	6 átomos, com os átomos e centroides relacionados por bases.	
<b>Tabela 3.27</b>	Demonstrativo das interações $\pi$ com origem na base G14, anel aromático com 5 átomos, com os átomos e centroides relacionados por bases.	106
<b>Tabela 3.28</b>	Demonstrativo das interações $\pi$ com origem na base C15, com os átomos e centroides relacionados por bases.	106
<b>Tabela 3.29</b>	Demonstrativo das interações $\pi$ com origem na base G16, anel aromático com 6 átomos, com os átomos e centroides relacionados por bases.	106
<b>Tabela 3.30</b>	Demonstrativo das interações $\pi$ com origem na base G16, anel aromático com 5 átomos, com os átomos e centroides relacionados por bases.	107
<b>Tabela 3.31</b>	Demonstrativo das interações $\pi$ com origem na base A17, anel aromático com 6 átomos, com os átomos e centroides relacionados por bases.	107
<b>Tabela 3.32</b>	Demonstrativo das interações $\pi$ com origem na base A17, anel aromático com 5 átomos, com os átomos e centroides relacionados por bases.	108
<b>Tabela 3.33</b>	Demonstrativo das interações $\pi$ com origem na base A18, anel aromático com 6 átomos, com os átomos e centroides relacionados por bases.	108
<b>Tabela 3.34</b>	Demonstrativo das interações $\pi$ com origem na base A18, anel aromático com 5 átomos, com os átomos e centroides relacionados por bases.	109
<b>Tabela 3.35</b>	Demonstrativo das interações $\pi$ com origem na base T19, com os átomos e centroides relacionados por bases.	109
<b>Tabela 3.36</b>	Demonstrativo das interações $\pi$ com origem na base T20, com os átomos e centroides relacionados por bases.	110
<b>Tabela 3.37</b>	Demonstrativo das interações $\pi$ com origem na base C21, com os átomos e centroides relacionados por bases.	110
<b>Tabela 3.38</b>	Demonstrativo das interações $\pi$ com origem na base G22, anel aromático com 6 átomos, com os átomos e centroides relacionados por bases.	111
<b>Tabela 3.39</b>	Demonstrativo das interações $\pi$ com origem na base G22, anel aromático com 6 átomos, com os átomos e centroides relacionados por bases.	111
<b>Tabela 3.40</b>	Demonstrativo das interações $\pi$ com origem na base C23, com os átomos e centroides relacionados por bases.	112
<b>Tabela 3.41</b>	Demonstrativo das interações $\pi$ com origem na base G24, anel aromático com 6 átomos, com os átomos e centroides relacionados por bases.	112
<b>Tabela 3.42</b>	Demonstrativo das interações $\pi$ com origem na base G24, anel aromático com 6 átomos, com os átomos e centroides relacionados por bases.	112
<b>Tabela 3.43</b>	Valor do deslocamento dos centroides de estrutura de DNA com ligante no	115



sulco, tendo como referência a posição do mesmo centroide em uma estrutura sem ligante.

<b>Tabela 3.44</b>	Valor do deslocamento dos centroides de estrutura de DNA com intercalado, tendo como referência a posição do mesmo centroide em uma estrutura sem ligante.	117
<b>Tabela 3.45</b>	Comparativo das interações $N^{\delta-}\pi$ identificadas por estrutura de DNA, não computadas as interações presentes nos ligantes.	121
<b>Tabela 3.46</b>	Comparativo das interações $O^{\delta-}\pi$ identificadas por estrutura de DNA, não computadas as interações presentes nos ligantes.	122
<b>Tabela 3.47</b>	Comparativo das interações $\pi^{\delta-}\pi$ identificadas por estrutura de DNA, não computadas as interações presentes nos ligantes.	124
<b>Tabela 3.48</b>	Comparativo das interações $C-H^{\delta-}\pi$ identificadas por estrutura de DNA, não computadas as interações presentes nos ligantes.	126
<b>Tabela 3.49</b>	Comparativo das interações $N-H^{\delta-}\pi$ identificadas por estrutura de DNA, não computadas as interações presentes nos ligantes.	128
<b>Tabela 3.50</b>	Comparativo entre os valores teóricos e os obtidos pelo mapeamento vetorial para os parâmetros geométricos para as interações $\pi$ , por tipo das estruturas de DNA estudadas.	136

## LISTA DE ABREVIATURAS E SIGLAS

°	<i>Grau</i>
Å	<i>Ångström</i>
BioMat	<i>Grupo de Física Teórica de Materiais e Biomoléculas</i>
CIF	<i>Crystallographic Information File</i>
DNA	<i>Deoxyribonucleic acid (ácido desoxirribonucléico)</i>
EBI	<i>European Bioinformatics Institute</i>
EDSAC	<i>Electronic Delay Storage Automatic Calculator</i>
FORTTRAN	<i>Formula Translator</i>
HB	<i>Hydrogen Bond</i>
IBM	<i>Industry of Business Machines</i>
IUCr	<i>International Union of Crystallography</i>
LaCrEMM	<i>Laboratório de Cristalografia Estereodinâmica e Modelagem Molecular</i>
MDB	<i>Macromolecules Data Base</i>
PDB	<i>Protein Data Bank</i>
RSCB	<i>Research Collaboratory for Structure Bioinformatics</i>

### Lista de nucleotídeos:

C	Citosina	A	Adenina
T	Timina	G	Guanina

### Lista de aminoácidos utilizados:

HIS	Histidina	TRP	Triptofano
PHE	Fenilalanina	TYR	Tirosina

## LISTA DE ANEXOS

<b>Anexo 3.1</b>	<i>1GRA MOL2 - MAPEAMENTO DE INTERAÇÕES <math>\pi</math></i>
<b>Anexo 3.2</b>	<i>DNA-SL 1BNA - MOL2 - H PROTEINAS ATOMOS-CENTROIDES</i>
<b>Anexo 3.3</b>	<i>DNA-SL 1BNA - MOL2 - ATOMOS-CENTROIDES</i>
<b>Anexo 3.4</b>	<i>DNA-LS 1VZK - MOL2 - H PROTEINAS ATOMOS-CENTROIDES</i>
<b>Anexo 3.5</b>	<i>DNA-LS 1VZK - MOL2 - ATOMOS-CENTROIDES</i>
<b>Anexo 3.6</b>	<i>DNA-LI 1G3X - MOL2 - H PROTEINAS ATOMOS-CENTROIDES</i>
<b>Anexo 3.7</b>	<i>DNA-LI 1G3X - MOL2 - ATOMOS-CENTROIDES</i>
<b>Anexo 3.8</b>	<i>Tabulação e Estatística DNA</i>
<b>Anexo 3.9</b>	<i>Estruturas de DNA com interações <math>\pi</math></i>
<b>Anexo 3.10</b>	<i>Estruturas de DNA com interações <math>\pi</math> sem as do ligante</i>
<b>Anexo 3.11</b>	<i>Estruturas de DNA com interações <math>\pi</math> por tipo com ligantes</i>
<b>Anexo 3.12</b>	<i>Estruturas de DNA com interações <math>\pi</math> por tipo sem ligantes</i>
<b>Anexo 3.13</b>	<i>Deslocamento das coordenadas dos centroides - base DNA sem ligante</i>
<b>Anexo 3.14</b>	<i>Estruturas de aminoácidos e bases do DNA com a identificação dos átomos</i>

**Observação:** Os anexos encontram-se em mídia digital ao final deste trabalho.

# SUMÁRIO

Resumo

Abstract

Lista de Figuras

Lista de Tabelas

Lista de Abreviaturas e Siglas

Lista de Anexos

## Apresentação

### Capítulo 1 - Introdução

página

1.1	Química Supramolecular	22
1.2	As Interações $\pi$	25
1.2.1	Forças Eletrostáticas	25
1.2.2	Interações $\pi$ ... $\pi$	29
1.2.3	Interações $\pi$ em Forma de "T"	33
1.2.4	Interações $\pi$ com o Hidrogênio – X-H... $\pi$	35
1.2.5	Interações cátion... $\pi$	37
1.3	Cristalografia e Química Supramolecular	40
1.4	Bancos de Dados	42
1.5	Ciência da Computação e Química Supramolecular	45
1.6	Visualizadores Gráficos	59
1.7	Conclusões do Capítulo 1	66

### Capítulo 2 – Materiais e métodos

página

2.1	Álgebra Vetorial e Química Supramolecular	68
2.1.1	O Conceito Vetorial Aplicado às Estruturas Cristalográficas	72
2.1.2	Cálculo da Coordenada do Centroide	73
2.1.3	Cálculo do Vetor Distância e seu Módulo	74
2.1.4	Cálculo do Vetor Normal ao Plano do Anel Aromático	75
2.1.5	Cálculo do Ângulo entre o Vetor Normal e o Vetor Distância	76

2.2	Utilização de Planilha Eletrônica para Cálculo Vetorial	77
2.3	Cálculos Utilizando Planilha Eletrônica	80
2.4	Conclusões do Capítulo 2	88
<b>Capítulo 3 – Resultados e Discussões</b>		<b>página</b>
3.1	Análises Qualitativas e Quantitativas de Interações $\pi$	92
3.1.1	Variação das Interações $\pi$ em Função da Presença de Ligantes	92
3.1.2	Mapeamento das Variações das Interações $\pi$	97
3.1.3	Interações $\pi$ Específicas nos Ligantes	113
3.2	Deformação do DNA em Função da Presença de Ligantes	114
3.3	Caminhos Eletrostáticos das Interações $\pi$	119
3.4	Estatística dos Parâmetros Geométricos das Interações $\pi$	130
3.5	Conclusões do Capítulo 3	136
<b>Capítulo 4 – Conclusões</b>		<b>página</b>
	Conclusões	139
<b>Referências</b>		<b>página</b>
	Referências	143

# Capítulo 1

---

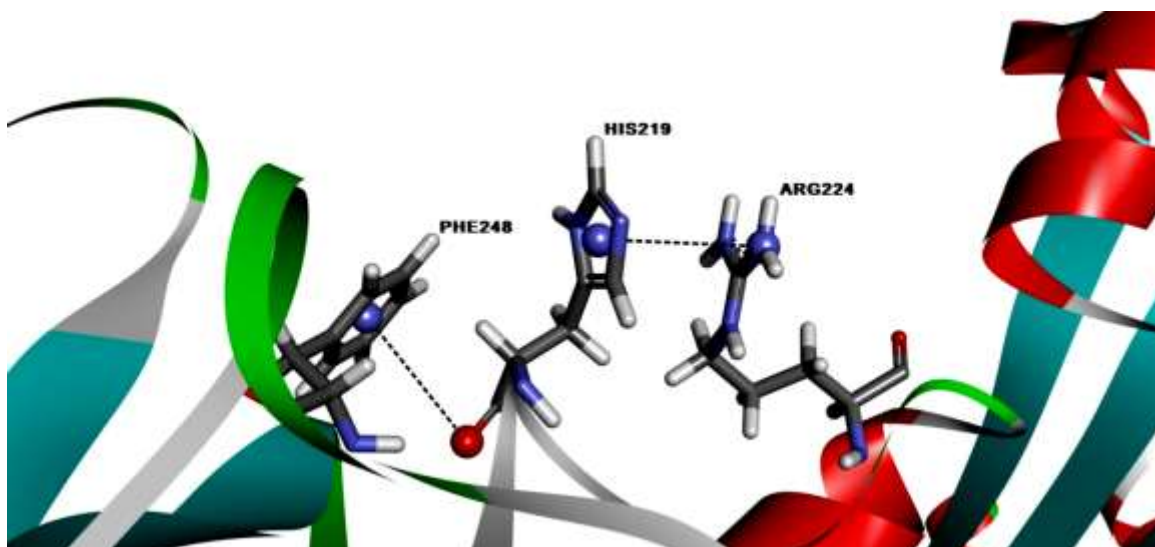
INTRODUÇÃO

O objetivo deste capítulo é proporcionar uma visão multidisciplinar coerente com a pesquisa desenvolvida. Há uma abordagem sobre os conceitos inerentes às interações  $\pi$ , mostrando a sua relevância para os sistemas biológicos, sua tipologia e características específicas, bem como os parâmetros geométricos que são essenciais para sua descrição. Discorre-se brevemente sobre os ensaios cristalográficos, os quais segundo a pesquisa efetuada mostraram ser a principal fonte de obtenção das estruturas depositadas e disponibilizadas em bancos de dados. Por fim, uma retrospectiva histórica para que pudesse ser observada a evolução da Ciência da Computação associada aos estudos envolvendo macromoléculas, finalizando com softwares específicos para a visualização “*in silico*” de estruturas biológicas que estão disponíveis em bancos de dados.

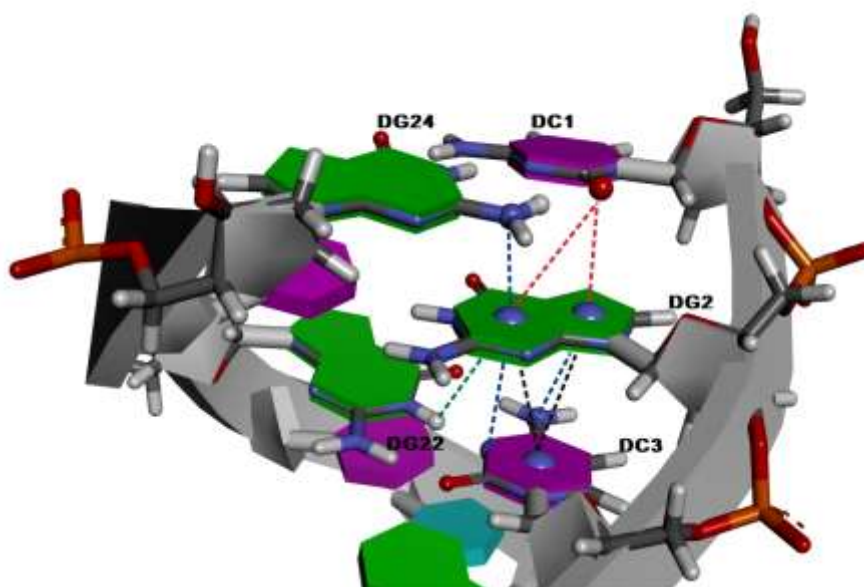
## 1.1 QUÍMICA SUPRAMOLECULAR

A química supramolecular é definida como a química dos arranjos moleculares e das ligações intermoleculares e está interessada em entidades de grande complexidade que resultam da associação de duas ou mais espécies químicas unidas por forças não-covalentes intermoleculares, Lehn (1995). Estes tipos de interações têm sido estudadas em pequenas moléculas desde 1995 por Zukerman-Schpector e colaboradores no Laboratório de Cristalografia, Estereodinâmica e Modelagem Molecular (LaCrEMM) no Departamento de Química da Universidade Federal de São Carlos (LaCrEMM, UFSCar) .

Entre as interações não-covalentes de grande importância estão as que envolvem sistemas delocalizados  $\pi$ , conhecidas simplesmente como interações  $\pi$ , que têm merecido destaque, principalmente por seu papel na engenharia de cristais como apresentado no trabalho de Zukerman-Schpector e Haiduc (2002). A sua análise e entendimento tem sido essenciais para descrever também o enovelamento de proteínas (Figura 1.1) e interações de empilhamento do DNA (Figura 1.2), ver por exemplo Seng e colaboradores (2010); Ermondi e Caron (2006); Mignon et al. (2005); Tewari e Dubey (2008).



**Figura 1.1** Representação gráfica de parte do enovelamento de uma glutatona redutase com identificação PDB 1GRA, onde podem ser observadas as interações  $O^{\pi} \pi$  entre a PHE248 e a HIS219 e a  $N^{\pi} \pi$  entre a HIS219 e a ARG224. Essas interações contribuem para a conformação do folding da proteína. Estrutura gerada a partir de dados cristalográficos, Karplus e Schulz (1989), utilizando-se o software Discovery Studio Client 3.1<sup>TM</sup>.



**Figura 1.2** Representação gráfica de parte de uma molécula de DNA identificada no PDB como 1BNA, com as interações  $\pi$  existentes entre as bases. Estrutura gerada a partir de dados cristalográficos por Drew et al. (1981) e utilizando-se o software Discovery Studio Client 3.1<sup>TM</sup>.

As interações  $\pi$  têm sido descritas em trabalhos de pesquisa realizados no Grupo de Física Teórica de Materiais e Biomoléculas (BioMat) no Departamento de Física



da Universidade Federal de São Carlos e tem mostrado serem fundamentais para entender as interações receptor-fármaco, tanto no caso de proteínas (Caracelli, I.; Zukerman-Schpector, J.; Maganhi, S. H.; Stefani, H. A.; Guadagnin, R.; Tiekink, E. R. T., 2010; Caracelli, I.; Zukerman-Schpector, J.; Tiekink, E. R. T., 2012; Vega-Teijido, M.; Caracelli, I.; Zukerman-Schpector, J., 2006) como de DNA (Caracelli, Zukerman-Schpector, Tiekink, 2012; Seng, H.-L.; Von, S.-T.; Tan, K.-W.; Maah, M.; Ng, S.-W.; R., Raja; Caracelli, I.; Ng, C.-H.; 2010).

A configuração típica de uma interação  $\pi$  pode ser visualizada na Figura 1.3, onde são apresentados os parâmetros que foram utilizados na caracterização de uma interação  $\pi$ . Eles são:

**C<sub>g</sub>** : centroide do anel;

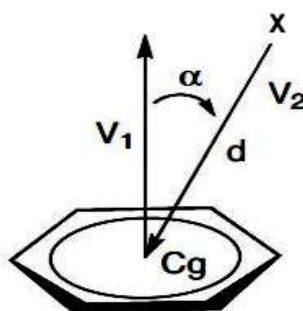
**X** : átomo ou centroide envolvido na interação  $\pi$  ;

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

Basicamente uma interação  $\pi$  existe em função de anéis aromáticos e/ou heteroaromáticos do ligante interagindo com as cadeias laterais de aminoácidos ou seus anéis; Meyer et al. (2003); Tiekink e Zukerman-Schpector (2010).

É possível evidenciar a relevância desse tipo de interações pelo fato de que as interações  $\pi$ - $\pi$  e as  $\pi$ -arila estão presentes em 20% dos fármacos. Nos sistemas biológicos, as interações  $\pi$  de maior relevância são as  $\pi$ - $\pi$ , O-H $\pi$ , N-H $\pi$  e cátion $\pi$ , Tewari (Tewari e Dubey, 2008). Também podem ser incluídas nessa modalidade as interações O $\pi$  e N $\pi$ , conforme Imai e colaboradores (2009), sendo que as características dessas interações são dependentes tanto das propriedades dos ligantes quanto dos sistemas  $\pi$ , Ermondi (Ermondi e Caron, 2006).

A ocorrência dessas interações, em proteínas dá-se pela presença dos aminoácidos: Fenilalanina (Phe), Histidina (His), Triptofano (Trp) e Tyrosina (Tyr), das bases do DNA: Adenina (DA), Citosina (DC), Glutamina (DG) e Timina (DT) e ainda quando há ligantes que possuam anéis aromáticos, McGaughey et al. (1998).

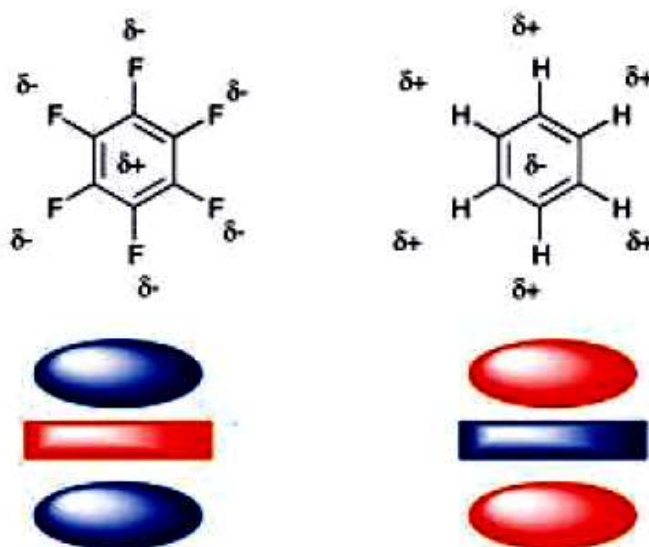
Enquanto que nas ligações covalentes as distâncias envolvidas são da ordem de 2,0 Å ou menos, interações  $\pi$  possuem uma variação maior de comprimento em função dos elementos envolvidos (Kiviniemi, 2001; Babu, 2003), sendo que o ângulo  $\alpha$  existente entre  $V_1$  e  $V_2$  considera-se em linhas gerais, menor ou igual a 30° (Babu, 2003; Caracelli, Zukerman-Schpector e Tiekink, 2012). Estas interações por possuírem a origem no centroide do plano do anel também são caracterizadas por sua direcionalidade.

## 1.2 AS INTERAÇÕES $\pi$

### 1.2.1 FORÇAS ELETROSTÁTICAS

São definidas por Interações  $\pi$  as interações não covalentes envolvendo anéis aromáticos, sendo essenciais para a estabilização das proteínas, DNA, moléculas orgânicas e nanomateriais, Smith et al. (2008). Essas interações estão presentes em moléculas que

possuem em sua conformação anéis aromáticos, possuindo estes uma distribuição de cargas eletrostáticas, onde a nuvem eletrônica existente forma um dipolo em cada lado do anel (Figura 1.4), dando origem a uma interação eletrostática (Kiviniemi, 2001).



**Figura 1.4** Representação da distribuição de cargas do hexafluoreno e benzeno. Essas estruturas moleculares apresentam dois dipolos opostos. Adaptado de Hunter et al. (2001).

A força das interações  $\pi$  é determinada pela combinação entre forças de atração (dispersivas, eletrostática e indutiva) e de repulsão (repulsão de troca), Singh et al. (2009). Considerando-se o balanço de energia em interações intermoleculares não covalentes descritas por Hunter e Sanders (1990), pode-se relacionar a energia total envolvida na interação entre duas moléculas pela Equação 1.1:

$$E_{total} = E_{eletrostática} + E_{indução} + E_{dispersão} + E_{repulsão} \quad (1.1)$$

De acordo com Ermondi e Caron (2006), para compreensão da Equação (1.1), faz-se necessária uma análise das forças que a compõem. Iniciando-se pelas forças

eletrostáticas que dependem da presença de cargas e/ou dipolos, é necessário conhecer os conceitos referentes ao momento de dipolo ( $\vec{p}$ ), polarizabilidade ( $\alpha$ ) e dipolo induzido ( $P_{ind}$ ).

O momento de dipolo  $\vec{p}$  entre duas cargas iguais de sinais diferentes e separadas por uma distância  $\vec{r}$  é um produto vetorial entre o valor da carga  $q$  e a distância  $\vec{r}$ .

$$\vec{p} = q \times \vec{r} \quad (1.2)$$

Já a polarizabilidade ( $\alpha$ ) é a facilidade com que a nuvem eletrônica de uma molécula sofre deformação quando exposta à um campo eletrostático e é mensurada pela razão entre o dipolo induzido ( $P_{ind}$ ) e o campo elétrico ( $E'$ ) que o induz. Com relação ao campo elétrico, assume-se como a média dos campos elétricos resultante das direções  $x$ ,  $y$  e  $z$ .

$$\alpha = \frac{P_{ind}}{E'} \quad (1.3)$$

As forças eletrostáticas subdividem-se em forças eletrostáticas simples e forças eletrostáticas combinadas.

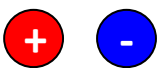

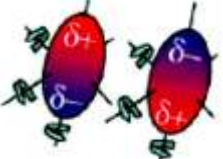


As forças eletrostáticas simples são resultantes das propriedades da partícula e também de fatores geométricos envolvendo a posição das partículas no sistema.

As forças eletrostáticas combinadas envolvem as forças eletrostáticas simples e fenômenos quânticos. Exemplos deste tipo de interação são as ligações de Hidrogênio e nos sistemas  $\pi$  (Tabela 1.1).

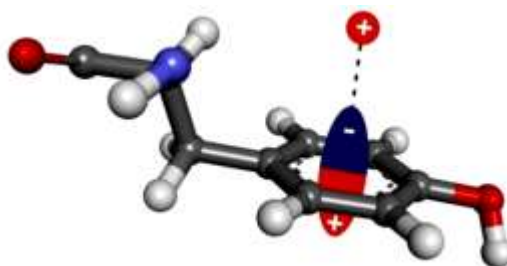
Os sistemas  $\pi$ , além da posição relativa dos anéis na interação, sofrem influência dos átomos ou heteroátomos próximos (Ermondi e Caron, 2006), uma vez que a atração eletrostática exercida por átomos carregados positiva ou negativamente ou o alinhamento de dipolos, polarizam a nuvem eletrônica do anel formando um dipolo

induzido, o que pode ser observado na Figura 1.5; (Hunter e Sanders, 1990; Hunter, Lawson et al., 2001).

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

Categoria	Representação esquemática	Equação da energia potencial	Energia aproximada por interação
Ion – Ion		$(4\pi\epsilon_0) \cdot r$	25 kJ.mol <sup>-1</sup>
Ion – Dipolo permanente		$-\frac{q^2 p^2}{(4\pi\epsilon_0)} \cdot \frac{1}{kTr^4}$	50 a 200 kJ.mol <sup>-1</sup>
Dipolo permanente – Dipolo permanente		$-\frac{2}{3} \cdot \frac{p_1^2 p_2^2}{(4\pi\epsilon_0)^2} \cdot \frac{1}{kTr^6}$	50 a 500 kJ.mol <sup>-1</sup>
Dipolo permanente – Dipolo induzido		$-\frac{p^2 \alpha_p}{(4\pi\epsilon_0)^2} \cdot \frac{1}{r^6}$	< 5 kJ.mol <sup>-1</sup>
Dipolo induzido-Dipolo induzido (Forças de dispersão de London)		Sempre presentes. Possuem a sua origem na mecânica quântica, são atrativas e variam na proporção e 1/r <sup>6</sup> .	< 5 kJ.mol <sup>-1</sup>
Forças de repulsão	Sempre presentes e possuem a sua origem na mecânica quântica. A repulsão é consequência da sobreposição das nuvens eletrônicas.		

Notas:  $q$  = carga do íon;  $\pi$  = momento do dipolo;  $\alpha_p$  = polarizabilidade;  $r$  = distância entre as partículas;  $\epsilon_0$  = permissividade no vácuo = 8,85x10<sup>-12</sup>C/N.m<sup>2</sup>;  $k$  = constante de Boltzmann;  $T$  = temperatura.



**Figura 1.5** 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 Client 3.1™.

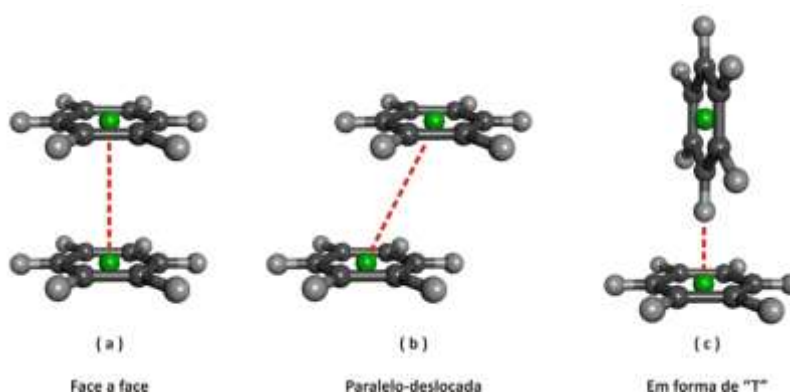
As forças eletrostáticas combinadas, conforme o seu tipo e características possuem diferentes níveis de energia por interação (Tabela 1.2).

**Tabela 1.2** Forças eletrostáticas combinadas, adaptado de Ermondi e Caron (2006)

Categoria	Classificação e energia aproximada por interação	
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>

### 1.2.2 INTERAÇÕES $\pi\cdots\pi$

A literatura refere-se a três tipos de interações  $\pi$  (Figura 1.6) que ocorrem nos anéis aromáticos cujos átomos estão com hibridização  $sp^2$  (Discovery Studio Client 3.1™, 2012), as interações  $\pi$  face a face, interações  $\pi$  paralelas-deslocadas, as interações  $\pi$  em forma de “T”, Meyer, Castelano et al. (2003).

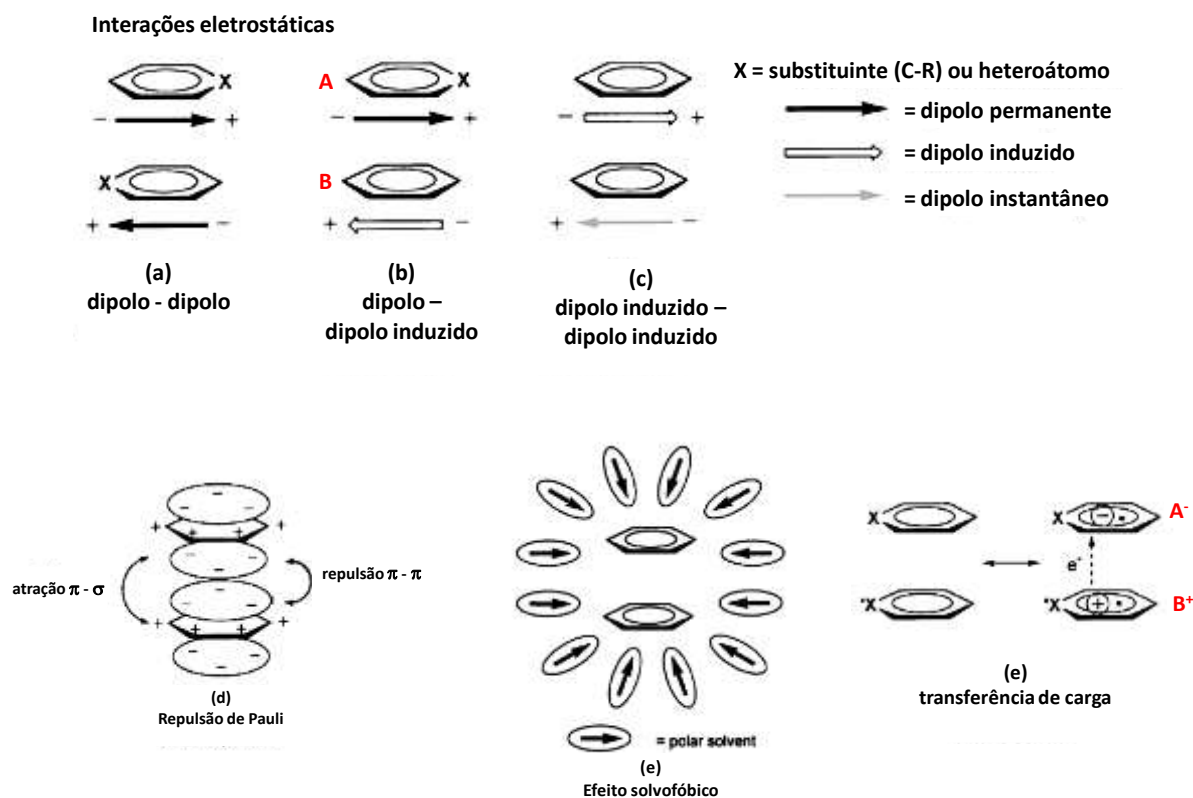


**Figura 1.6** Representações gráficas dos tipos de interação  $\pi\cdots\pi$ : (a) face a face, (b) paralela-deslocada, (c) forma de “T”. Adaptado de Singh et al. (2009) e Sherril et al. (2009).

Conforme Burley e Petsko (1985), em média 60% dos aminoácidos aromáticos presentes nas cadeias laterais de proteínas estão envolvidos em interações  $\pi\cdots\pi$ , e as do

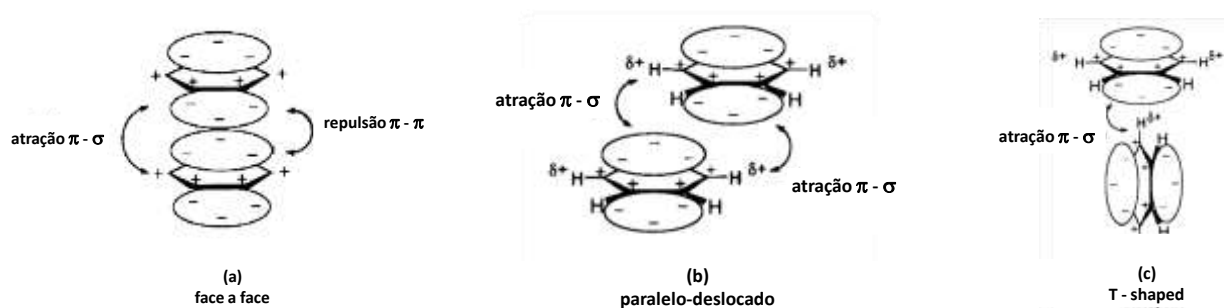
tipo T-shaped são maioria. Já McGaughey e colaboradores (Mcgaughey et al., 1998), em um estudo mais amplo, sugere a geometria paralelo-deslocada como a configuração mais comum. Meyer e colaboradores (Meyer, Castellano e Diederich, 2003), em seu review, observa que outros estudos em estruturas do PDB concluem que as geometrias T-shaped e paralelo-deslocada são as predominantes, não sendo encontrados estudos sobre o empilhamento face a face (Janiak, 2000). São determinantes na sua ocorrência a posição do anel aromático na proteína, os tipos de anéis envolvidos e suas distâncias.

As interações  $\pi\cdots\pi$  são caracterizadas pela combinação de forças de atração entre a estrutura  $\sigma$  de um anel com a componente  $\pi$  e a repulsão das nuvens eletrônicas dos anéis que compõem o sistema; sendo que a força  $\sigma\cdots\pi$  possui maior intensidade que a repulsão  $\pi\cdots\pi$  (Hunter e Sanders, 1990; Hunter et al., 2001). Conforme Ermondi e Caron (2006), por se tratarem de forças combinadas, existem também interações entre os dipolos. Janiak (2000) relaciona os tipos de interações envolvidos nos sistemas  $\pi\cdots\pi$ , sendo elas: eletrostáticas, de repulsão e transferência de carga (Figura 1.7).



**Figura 1.7** Interações  $\pi\cdots\pi$ . (a) interações dipolo-dipolo: interações entre as distribuições de carga permanentes e estáticas; (b) interações dipolo-dipolo induzido: interações entre a carga estática do grupo A e carga induzida pela proximidade no grupo B; (c) interações dipolo induzido-dipolo induzido: o momento de dipolo instantâneo de uma nuvem de elétrons polariza uma molécula vizinha e induz nele também um dipolo instantâneo; (d) Força de Repulsão de Pauli: a curtas distâncias as nuvens eletrônicas geram uma força de repulsão  $\pi\cdots\pi$  que torna-se superior à força de atração  $\pi\cdots\sigma$ ; (e) Efeitos solvofóbicos: em solventes polares, efeitos solvofóbicos que conduzem a dessolvatação podem estabilizar uma agregação de superfícies moleculares lipofílicas; (f) transferência de carga: o efeito de transferência de carga ou elétron doador-aceptor entre o grupo A<sup>-</sup> e o B<sup>+</sup> criam uma força de atração. Adaptado de Janiak (2000).

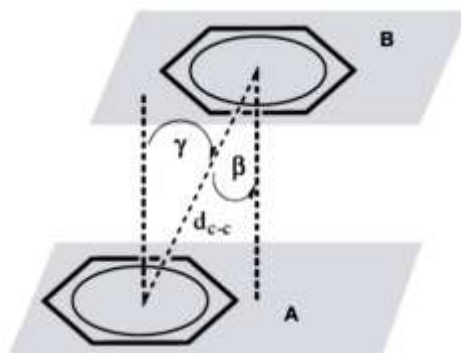
Assim sendo, a geometria das interações  $\pi\cdots\pi$  está relacionada com a possível combinação de duas ou mais forças decorrentes das interações das nuvens eletrônicas entre si e com a estrutura  $\sigma$  dos anéis ou campos elétricos próximos a estes (Burley e Petsko, 1985). A relação entre as forças de atração  $\pi\cdots\sigma$  e repulsão  $\pi\cdots\pi$  em um arranjo espacial é alterada em função do posicionamento entre os anéis, justificando assim as configurações paralelo-deslocada e em forma de “T” como interações possíveis, sendo a interação  $\pi\cdots\pi$  face a face um modelo teórico, Meyer et al. (2003) e Janiak (2000). O esquema das cargas para as três configurações encontra-se na Figura 1.8.



**Figura 1.8** Interações  $\pi\cdots\pi$ , esquema de distribuição de cargas. (a) face a face, (b) paralela-deslocada, (c) em forma de “T”. Adaptado de Janiak (Janiak, 2000).

Dessa forma, para estudo de uma interação  $\pi$  face a face e face-deslocada, são levados em consideração os parâmetros geométricos em que se encontram os anéis aromáticos (Figura 1.9), (Meyer et al., 2003); Janiak, 2000; Tomić et al., 2006).





**Figura 1.9** Parâmetros geométricos para as interações  $\pi\cdots\pi$ . (Adaptado de Tomić et al., 2006).

Os parâmetros geométricos são:

$d_{c-c}$  : distância entre os centroides dos anéis aromáticos **A** e **B**;

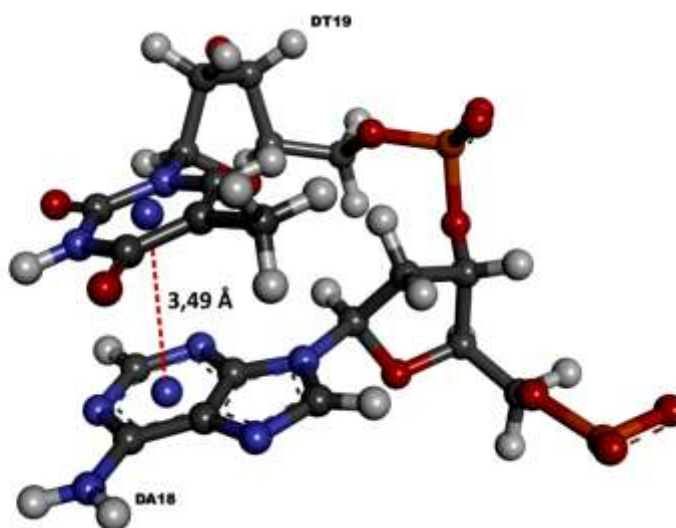
$\gamma$  : ângulo entre o vetor normal ao plano do anel **A** e  $d_{c-c}$ ;

$\beta$  : ângulo entre o vetor normal ao plano do anel **B** e  $d_{c-c}$ ;

A bibliografia pesquisada versa diferentes valores para os parâmetros geométricos para as interações  $\pi$  face a face e face-deslocada, todavia, estes situam-se dentro de faixas específicas em função da natureza das moléculas envolvidas nos estudos.

A distância  $d_{c-c}$  está situada entre 3,0 e 4,6 Å, segundo Janiak (2000), sendo que em trabalhos de outros autores (Meyer et al., 2003; Tomić et al., 2006; Khavasi e Azizpoor Fard, 2010; Marsili et al., 2008; Seng et al., 2010), os valores especificados para a distância  $d_{c-c}$  estão contidos dentro dessa faixa. Com relação aos ângulos  $\beta$  e  $\gamma$ , há um consenso entre diversos autores de que o ângulo entre o vetor normal ao plano do anel aromático e o vetor  $d_{c-c}$  seja menor ou igual a 30° (Caracelli, Zukerman-Schpector e Tiekink, 2012; McGaughey et al., 1998; Babu, 2003; Tomić et al., 2006; Marsili et al., 2008; Zukerman- et al., 2011). Deve-se ainda considerar que para que efetivamente ocorra uma interação  $\pi\cdots\pi$ , os ângulos  $\beta$  e  $\gamma$  devem ambos atenderem ao requisito de serem iguais ou menores do que 30°. Para que seja caracterizada interação  $\pi\cdots\pi$  face a face, os ângulos  $\beta$  e  $\gamma$  devem obrigatoriamente serem iguais.

Um exemplo de uma interação  $\pi \cdots \pi$  em meio biológico é observado entre as bases de um DNA estudado nesse trabalho(Figura 1.10).

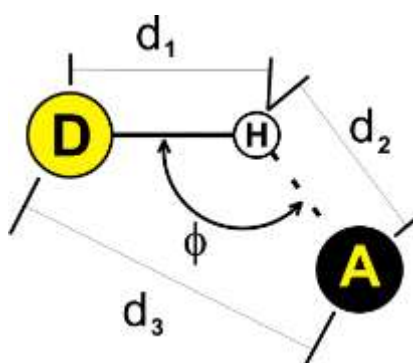


**Figura 1.10** Representação gráfica de uma interação  $\pi \cdots \pi$  entre as bases DA 18 e DT 19 de uma molécula de DNA com identificação PDB 1BNA. Estrutura gerada a partir de dados cristalográficos por Drew et al. (1981) com a utilização do software Discovery Studio Client 3.1™.

### 1.2.3 INTERAÇÕES $\pi$ EM FORMA DE “T”

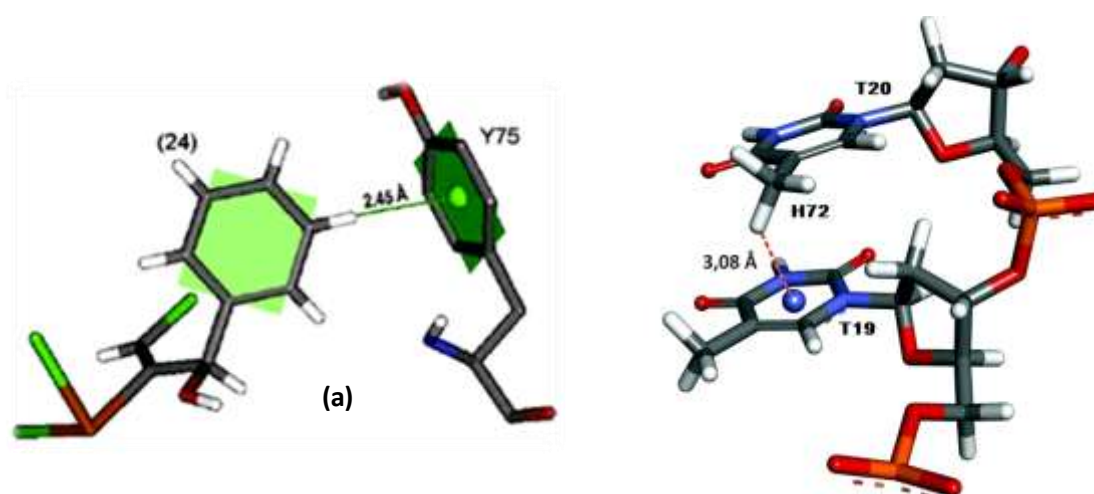
Embora a literatura aborde a configuração da interação  $\pi$  em forma de “T” conjuntamente com as interações  $\pi$  face a face e paralelo-deslocadas por esta envolver dois anéis aromáticos, trata-se efetivamente de uma interação  $\pi \cdots \sigma$ , onde apenas um dos centroides dos dois anéis aromáticos está interagindo com um átomo ligado covalentemente ao anel (Janiak, 2000). Este átomo em meios biológicos geralmente é o hidrogênio e sua função é a indução de um dipolo na nuvem eletrônica do anel aromático no qual está o centroide. Nessa interação, semelhantemente à uma ligação de hidrogênio (HB), a nuvem eletrônica do anel aromático (aceptor) interage com um átomo de hidrogênio ligado covalentemente à um segundo átomo eletronegativo pertencente ao segundo anel aromático (doador), sendo que essa interação é potencializada em função das dimensões do átomo de hidrogênio que permite a proximidade da interação entre dipolos e cargas mostrada na Figura 1.11 (Ermondi e Caron, 2006; Nishio, 2011; Shanthi et

al., 2010). Há ainda uma similaridade entre interação  $\pi$  em forma de “T” e a interação C-H $\cdots\pi$  (Cunha et al., 2006; Birchall et al., 2011), porém no caso desta última, o carbono ligado ao hidrogênio não pertence a um anel (Meyer et al., 2003).



**Figura 1.11** Representação esquemática de uma interação  $\pi\cdots\sigma$ , considerando como doador o átomo de um centroide ligado ao hidrogênio e aceptor a nuvem eletrônica do segundo anel aromático. (Adaptado de Ermondi e Caron, 2006).

Contudo, no caso da interação  $\pi$  em forma de “T” considera-se que a distância  $d_2$  possui como referência para sua mensuração a posição dos centroides dos dois anéis aromáticos. Para que seja possível obter o valor real do comprimento da interação da interação  $\pi$  em forma de “T”, deve-se subtrair a soma da distância do centroide do anel até o átomo ao qual está ligado covalentemente o hidrogênio e o comprimento da ligação de hidrogênio a este átomo. O total dessa soma no anel de 06 átomos de carbono em uma molécula de triptofano é de 2,5 Å. Assim, quando Tsuzuki e colaboradores (Tsuzuki et al., 2001) especifica que a distância entre os centroides na configuração T-shaped é de 5,0 Å, ao subtrair a distância total do centroide ao hidrogênio (2,5 Å), obtém-se a distância mínima de 2,5 Å para o comprimento desta interação. Por outro lado, para Babu (2003), o comprimento de uma interação C-H $\cdots\pi$  é igual ou menor a 3,8 Å. Dessa forma é possível definir a faixa do comprimento dessas interações entre 2,5 e 3,8 Å. Esses valores são corroborados pelas imagens abaixo (Figura 1.12).

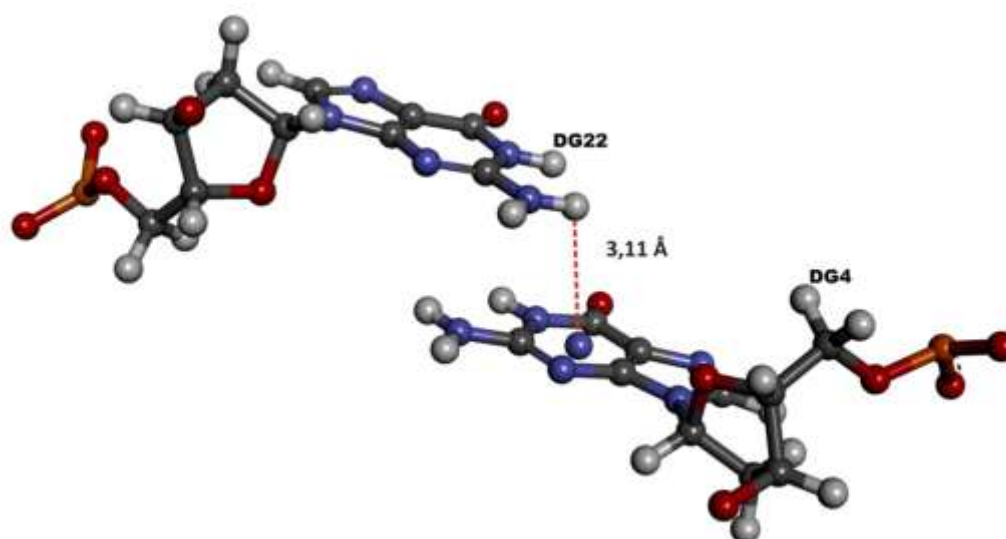


**Figura 1.12** (a) Representação gráfica de uma interação C-H... $\pi$  T-shaped, de Cunha et al. (2006); (b) Interação C-H... $\pi$  entre duas bases de uma molécula de DNA com identificação PDB 1BNA. Estrutura gerada a partir de dados cristalográficos por Drew e colaboradores (1981) com a utilização do software Discovery Studio Client 3.1<sup>TM</sup>.

Permanece o consenso de que o ângulo entre o vetor normal ao plano do anel aromático e o vetor  $d_2$  deve ser inferior ou igual a  $30^\circ$  (Caracelli, Zukerman-Schpector e Tiekink, 2012; Mcgaughey, Gagné, Rappé, 1998; Babu, 2003; Tomić et al., 2006; Marsili et al., 2008; Zukerman-Schpector et al., 2011).

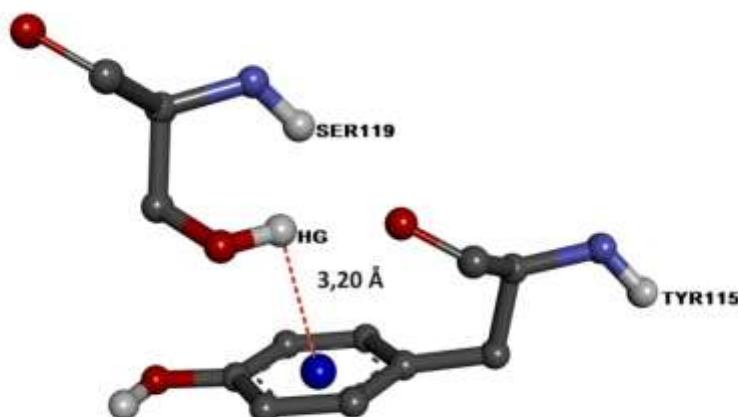
#### 1.2.4 INTERAÇÕES $\pi$ COM HIDROGÊNIO – X-H... $\pi$

Para que as interações X-H... $\pi$  ocorram, são necessários os mesmos parâmetros das interações C-H... $\pi$ , ou seja, a nuvem eletrônica do anel aromático (acceptor) interage com um átomo de hidrogênio de um grupo doador X-H (sendo X = O, N, S), (Sherrill et al., 2009; Mohan et al., 2010; Nishio, 2011; Tsuzuki et al., 2000). No grupo X-H o hidrogênio apresenta papel semelhante ao que ocorre nas ligações de hidrogênio, todavia, de acordo com Nishio (2011), esses tipos de interação são mais fracos que as ligações de hidrogênio. Juntamente com as interações C-H... $\pi$ , as interações N-H... $\pi$  e O-H... $\pi$  são responsáveis para estabilizar as proteínas e reconhecimento biológico (Mohan et al., 2010). Exemplos destas interações são ilustrados nas Figuras 1.13 e 1.14.



**Figura 1.13** Representação gráfica de uma interação N-H... $\pi$  entre duas bases de uma molécula de DNA com identificação PDB 1G3X. Estrutura gerada a partir de dados cristalográficos por Malinina et al. (2002) utilizando o software Discovery Studio Client 3.1<sup>TM</sup>.

Os parâmetros geométricos das interações X-H... $\pi$  segundo Babu (2003), são similares às interações C-H... $\pi$ , sendo o seu comprimento igual ou menor a 3,8 Å. Em seu estudo, Mohan e colaboradores (2010) apresenta valores entre 2,2 Å e 2,5 Å para interações e O-H... $\pi$  e 2,1 Å e 3,0 Å para as interações N-H... $\pi$ , estando estes valores compreendidos no limite estabelecido por Babu (2003).



**Figura 1.14** Representação gráfica de uma interação O-H... $\pi$  em uma cyclooxygenase-2 com identificação PDB 1CX2. Estrutura gerada a partir de dados cristalográficos por Kurumbail e colaboradores (2002) com a utilização do software Discovery Studio Client 3.1<sup>TM</sup>.

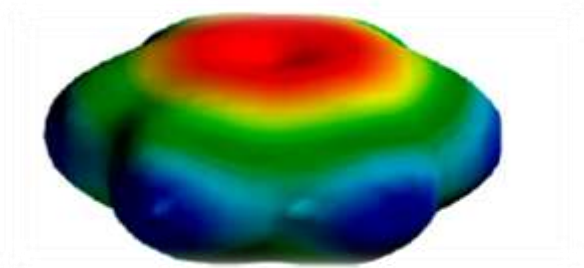
Ainda em seu estudo, Babu (2003) estipula que o ângulo entre o vetor normal ao plano do anel aromático e o vetor distância da interação deve ser menor ou igual a 30°, parâmetro este também adotado por outros pesquisadores (Caracelli, Zukerman-Schpector e Tiekink, 2012; Mcgaughey, Gagné, Rappé, 1998; Babu, 2003; Tomić et al., 2006; Marsili et al., 2008; Zukerman-Schpector, et al., 2011).

### 1.2.5 INTERAÇÕES CÁTION - $\pi$

As interações não covalentes do tipo Cátion $\cdots\pi$  ocorrem entre os anéis aromáticos (sistema  $\pi$ ) e cátions em sua adjacência, podendo estes cátions serem de origem inorgânica ou orgânica. 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) e todos os cátions metálicos, a exemplo do Potássio ( $K^+$ ), Sódio ( $Na^+$ ) e Cálcio ( $Ca^{2+}$ ), (Du et al., 2012; Dougherty (Ma e Dougherty, 1997; Dougherty, 2007).

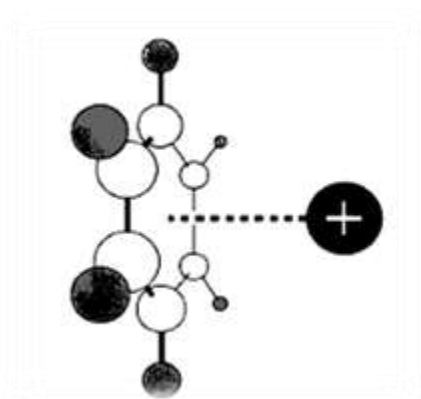
Em seus artigos, Zukerman-Schpector (Zukerman-Schpector e Haiduc, 2002; Tiekink e Zukerman-Schpector, 2009; Tiekink e Zukerman-Schpector, 2010; Zukerman-Schpector, J.; Otero-de-la-Roza, A.; Luana, V.; Tiekink, E., 2011), identificou esse tipo de interação em estruturas que contem os elementos Arsênio (As), Telúrio (Te), Chumbo (Pb) e Ouro (Au) e Caracelli (Caracelli, Zukerman-Schpector e Tiekink, 2012) identificou o Selênio (Se).

Essas interações são responsáveis pela estrutura secundária das proteínas e também na sua interação com ligantes e fármacos, estando atuando também na estabilização do DNA. Para que as interações Cátion $\cdots\pi$  ocorram, a nuvem eletrônica do anel aromático (Figura 1.15) atua como doadora de elétrons interagindo com a carga positiva do cátion (Du et al., 2012; Ma e Dougherty, 1997; Dougherty, 2007; Gromiha et al., 2004; Wintjens et al., 2000).



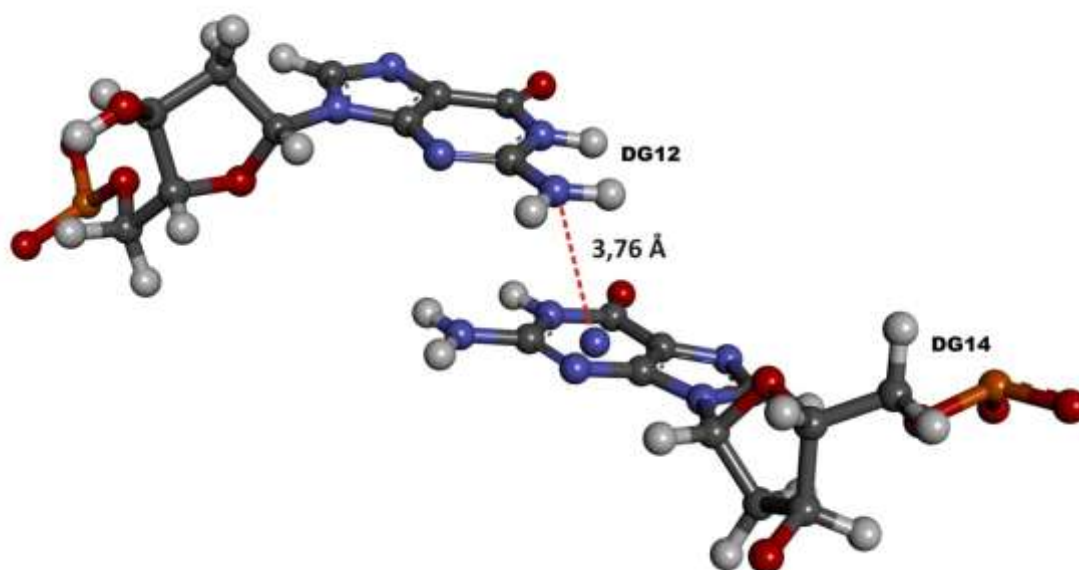
**Figura 1.15** Nuvem eletrônica de um anel aromático que atua como doadora de elétrons para uma interação cátion $\pi$ , sendo a região vermelha eletronegativa. (Extraído de Dougherty, 2007).

Ainda conforme Dougherty (2007), a geometria característica dessa interação é que o cátion deve estar posicionado sobre a face do anel aromático (Figura 1.16).

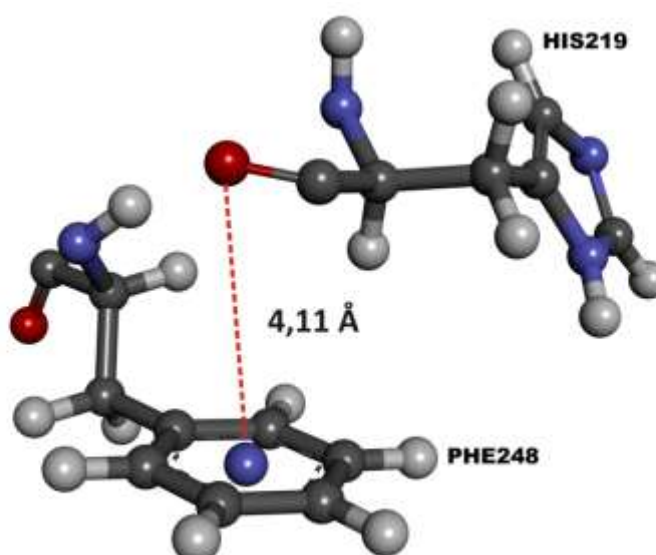


**Figura 1.16** Esquema de uma interação cátion $\pi$ , representando-se um cátion genérico posicionado sobre a face de um anel de benzeno. Extraído de Ma e Dougherty (1997).

Caracterizam-se como interações Cátion $\pi$  as interações N $\pi$  (Figura 1.17) e O $\pi$  (Figura 1.18), onde o nitrogênio e oxigênio são elétrons deficientes, segundo Zhang e colaboradores (2011).



**Figura 1.17** Representação gráfica de uma interação  $N \cdots \pi$  entre as bases G 12 e G 14 de uma molécula de DNA com identificação PDB 1BNA. Estrutura gerada a partir de dados cristalográficos por Drew et al. (1981) com a utilização do software Discovery Studio Client 3.1<sup>TM</sup>.



**Figura 1.18** Representação gráfica de uma interação  $O \cdots \pi$  entre a HIS219 e a PHE248 de glutatona reductase identificação PDB 1GRA. Estrutura gerada a partir de dados cristalográficos obtidos por (Karplus e Schulz (1989) visualizada com a utilização do software Discovery Studio Client 3.1<sup>TM</sup>.



Segundo a bibliografia pesquisada, os parâmetros geométricos para as interações Cation $\cdots\pi$  não possuem uma uniformidade com relação ao comprimento das interações e são escassas no tocante ao ângulo entre o vetor normal ao plano do anel aromático e o vetor distância da interação.

Com relação ao comprimento destas interações, Minoux e Chipot (1999) consideram que estas devem ser menores do que 5,0 Å para as interações entre fenilalanina e lisina, tirosina e lisina, triptofano e lisina e não fazem menção ao ângulo dessa interação com o vetor normal ao plano. Já Zhang e colaboradores (Zhang et al., 2011), estabelece uma faixa entre 3,6 Å e 3,8 Å para as interações nas estruturas por ele estudadas e um ângulo máximo de 5°. Sörme e colaboradores (2005) identificaram na estrutura de seu trabalho um comprimento de 4,3 Å, também não especificando o ângulo. Em seu artigo, Wouters (1998) especifica uma faixa de distância para a interação Cation $\cdots\pi$  entre 2,8 Å e 4,9 Å para cátions metálicos, sem também especificar os ângulos dessa interação com o vetor normal ao plano.

Zukerman-Schpector e colaboradores (Zukerman-Schpector e Haiduc, 2002; Tiekink e Zukerman-Schpector, 2009; Tiekink e Zukerman-Schpector, 2010; Zukerman-Schpector et al., 2011), adota que o comprimento das interações das estruturas por ele estudadas devem ser inferiores a 4,0 Å e o ângulo da interação com o vetor normal ao plano menor do que 20° e Caracelli (Caracelli, Zukerman-Schpector e Tiekink, 2012) também adota o valor de 4,0 Å, porém assumindo um ângulo máximo de 30° relativo entre o vetor distância e o normal ao plano.

### **1.3 CRISTALOGRAFIA E QUÍMICA SUPRAMOLECULAR**

Segundo Drenth (2007), as macromoléculas (proteínas, ácidos nucleicos, e hidratos de carbono) são os principais componentes não aquosos de células vivas, sendo que as proteínas proporcionalmente apresentam-se em maior quantidade (Tabela 1.3). Para que possam ser estudados os processos biológicos, o conhecimento da estrutura tridimensional das macromoléculas é de suma importância, conforme preconizado por Kendrew e colaboradores (1958).

**Tabela 1.3** Distribuição das estruturas de macromoléculas por tipo depositadas no PDB, adaptado de RSCB PDB

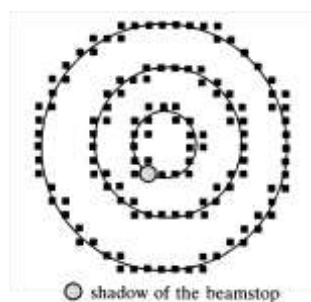
Tipos de Macromoléculas	Entradas em 09/09/2012	Proporção
Proteínas	78.110	92.6%
Ácidos Nucleicos	2.424	2.9%
Complexos Proteínas/Ácidos Nucleicos	3.823	4.5%
Outros	24	0.0%
<b>Total</b>	<b>84.831</b>	<b>100.0%</b>

A difração de raios X é a técnica mais utilizada para a obtenção de estruturas tridimensionais em alta resolução e também para macromoléculas de menor peso atômico (Tabela 1.4).

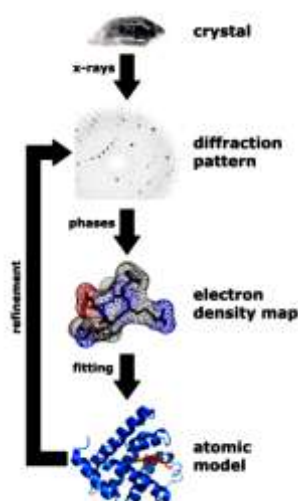
**Tabela 1.4** Métodos utilizados para a obtenção das estruturas de macromoléculas depositadas no PDB, adaptado de RSCB PDB

Método utilizado	Entradas em 09/09/2012	Proporção
Difração de raios X	74.131	87,9%
Ressonancia Magnetica Nuclear	9.583	11,4%
Microscopia eletrônica	453	0,5%
Híbrido	51	0,1%
Outros	163	0,2%
<b>Total</b>	<b>84.831</b>	<b>100,0%</b>

Essa técnica baseia-se na incidência de um feixe de raios X através de um cristal isolado de uma macromolécula (proteínas, ácidos nucleicos e hidratos de carbono) e a simetria dos átomos agrupados nesse cristal efetuará a dispersão em várias direções da radiação eletromagnética e por difração, tem-se um padrão (Figura 1.19) o qual resultará em um mapa de densidade eletrônica que submetido à softwares específicos fornece a estrutura tridimensional da macromolécula mostrado na Figura 1.20 (Drenth, 2007; Sun, 2004). Esta técnica possui uma alta complexidade na obtenção de um cristal livre de impurezas e na repetibilidade do processo (Caracelli e Zukerman-Schpector, 2009; Drenth, 2007).



**Figura 1.19** Representação esquemática do padrão de difração de um cristal de lisozima, extraído de Drenth (2007).



**Figura 1.20** Representação do processo de cristalografia por difração de raios X, extraído de Wikipedia (Contributors, 2012b).

Os resultados obtidos pela Cristalografia são depositados em centrais de coordenadas de estruturas cristalográficas de moléculas biológicas, constituindo importante fonte de pesquisa para a Química Supramolecular; Caracelli (Caracelli e Zukerman-Schpector, 2010).

#### 1.4 BANCOS DE DADOS

Uma vez reconhecida que a estrutura tridimensional de uma macromolécula é fundamental para o seu estudo (Kendrew et al., 1958) e que esta estrutura obtida principalmente pela cristalografia deve ser de fácil acesso aos usuários que necessitam destes dados para suas pesquisas, no início dos anos 70, Edgar Meyer, Gerson Cohen e

Helen Berman idealizaram a criação de uma central de coordenadas de estruturas cristalográficas de moléculas biológicas, Caracelli (Caracelli e Zukerman-Schpector, 2010). Dessa forma, em 1971 foi criado o Protein Data Bank (PDB) no Brookhaven National Laboratories com a função de banco de dados de estruturas de moléculas biológicas (Bernstein et al., 1977; Berman et al., 2000; Sussman et al., 1999; Sussman et al., 2012). Segundo Meyer (1997), que foi um dos idealizadores do PDB, o objetivo deste repositório de estruturas obtidas cristalograficamente é permitir que estas estruturas sejam processadas “*in silico*” e assim analisadas graficamente. Como todo banco de dados (Tenenbaum et al., 1995), o assim criado PDB apresentava uma estrutura, que veio a disponibilizar de forma uniforme os dados nele depositados, tais como coordenadas atômicas, citações, primárias e informação sobre a estrutura secundária, dados de estrutura experimentais, bem como hyperlinks para outros bancos de dados (Bernstein et al., 1977; Sussman et al., 2012).

A partir do final da década de 80, a expansão dos recursos de informática mostrou ser necessária a criação de um banco de dados mais estruturado para facilitar a publicação eletrônica e a partir de 1990, com a participação da União Internacional de Cristalografia (IUCr) iniciou-se um trabalho para desenvolver extensões macromoleculares para a representação de dados utilizado para a descrição das estruturas de pequenas moléculas e a determinação da estrutura cristalográfica, chamado Arquivo de Informações Cristalográficas (CIF) e passou a incluir definições detalhadas descrevendo a estrutura macromolecular e o experimento cristalográfico (Figura 1.21).

Em 1998 o Research Collaboratory for Structure Bioinformatics (RSCB) adotou o padrão criado como base do seu processamento de dados e infra-estrutura para a sua gestão. Além do RSCB, outras instituições ao redor do mundo adotaram o padrão CIF de forma a que a estrutura de dados viesse a tornar-se um padrão único. Essas instituições foram o European Bioinformatics Institute (Embl-Ebi, 1995) com o seu o Banco de Dados de Estruturas Macromoleculares (MDB) e o PDBj na Universidade de Osaka, formando estas o PDB (World Wide PDB – wwPDB), (Berman et al., 2007; Westbrook et al., 2005).

```

(a) Structured PDB records

CRYST1  63.150  83.590  53.800  90.00  99.34  90.00 P 21
ORIGX1   .963457   .136613   .230424   16.61000
ORIGX2  -.158977   .983924   .081383   13.72000
ORIGX3  -.215598  -.115048   .969683   37.65000
SCALE1   .015462   .002192   .003698   .26656
SCALE2  -.001902   .011771   .000974   .16413
SCALE3  -.001062  -.001721   .018728   .75059
ATOM     1  N  VAL A  1      6.130  16.559  4.905  7.00  41.29
ATOM     2  CA VAL A  1      6.870  17.784  4.702  6.00  41.33
ATOM     3  C  VAL A  1      8.377  17.548  4.913  6.00  31.64
ATOM     4  O  VAL A  1      8.820  16.980  5.922  8.00  38.31
ATOM     5  CB VAL A  1      6.345  18.763  5.731  6.00  52.26
ATOM     6  CG1 VAL A  1     6.745  20.188  5.356  6.00  52.75
ATOM     7  CG2 VAL A  1     4.826  18.612  5.847  6.00  58.75

(b) Unstructured PDB records

REMARK   3  REFINEMENT. BY THE METHOD OF JACK AND LEVITT. THE SCALE
REMARK   3  FACTOR BETWEEN ENERGY AND X-RAY FORMS WAS VARIED BETWEEN
REMARK   3  .00025 AND .0005 TO MAINTAIN THE RMS VARIATION OF THE C-C
REMARK   3  SINGLE BONDS BETWEEN 0.02 AND 0.03 ANGSTROMS. THE IRON
REMARK   3  ATOMS WERE UNRESTRAINED IN ORDER TO AVOID ANY POSSIBLE
REMARK   3  BIAS IN THEIR POSITIONS. THE FINAL R VALUE IS 0.16.

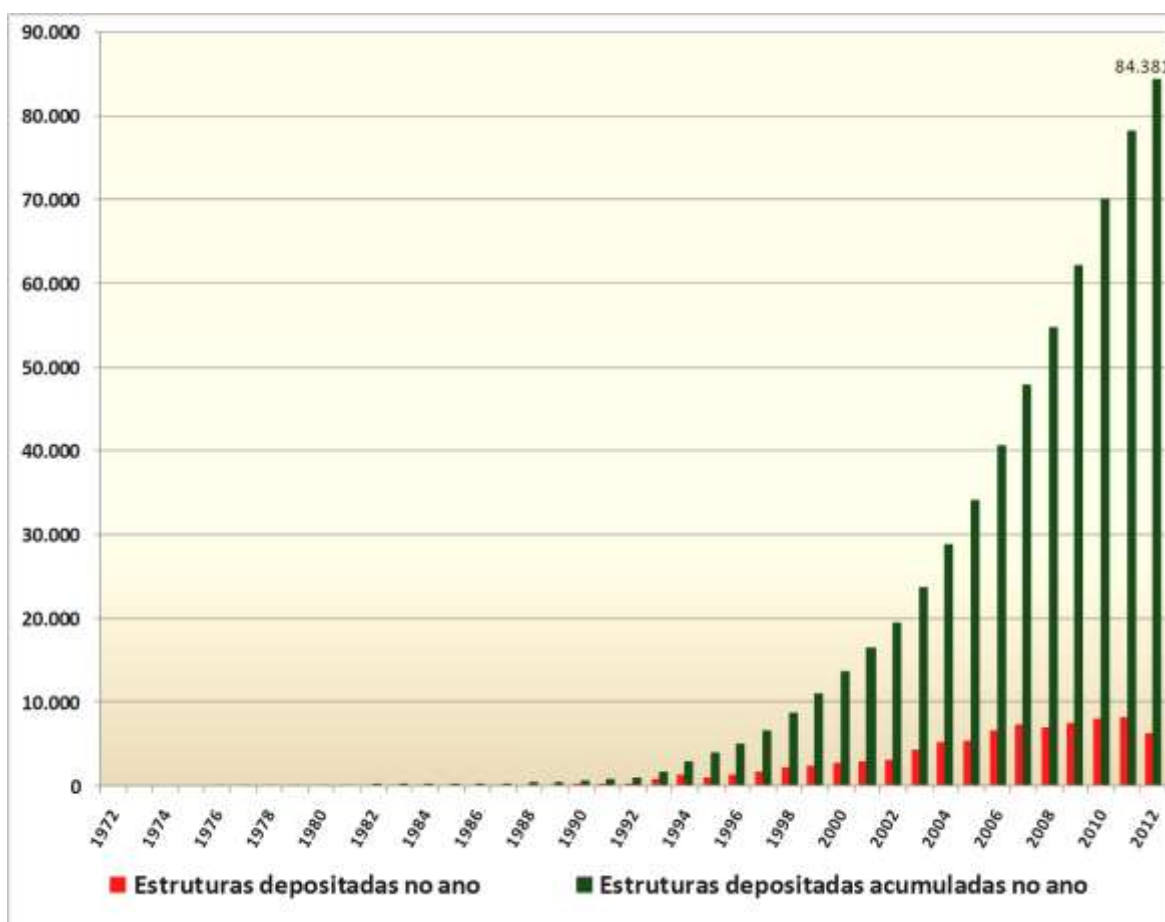
(c) Semi-structured PDB records

REMARK   3  FIT TO DATA USED IN REFINEMENT.
REMARK   3  CROSS-VALIDATION METHOD           : THROUGHOUT
REMARK   3  FREE R VALUE TEST SET SELECTION  : RANDOM
REMARK   3  R VALUE (WORKING SET)           : 0.213
REMARK   3  FREE R VALUE                     : 0.257
REMARK   3  FREE R VALUE TEST SET SIZE (%)   : 7.500
REMARK   3  FREE R VALUE TEST SET COUNT     : 2532
REMARK   3  ESTIMATED ERROR OF FREE R VALUE : 0.005

```

**Figura 1.21** Partes de registros de um arquivo de dados PDB. (a) registros estruturados do PDB descrevendo as constantes de uma cela cristalográfica (CRYST1), matrizes de transformação entre as coordenadas ortogonais e fracionada (ORIGX e escala) e as coordenadas atômicas (ATOM). (b) registros não-estruturados do PDB descrevendo os detalhes de refinamento cristalográfico usado em arquivos de dados do PDB antes de 1996. (c) registros semi-estruturados do PDB descrevendo o refinamento cristalográfico utilizado nos dados do PDB em arquivos após 1996. Extraído de Westbrook (Westbrook et al., 2005).

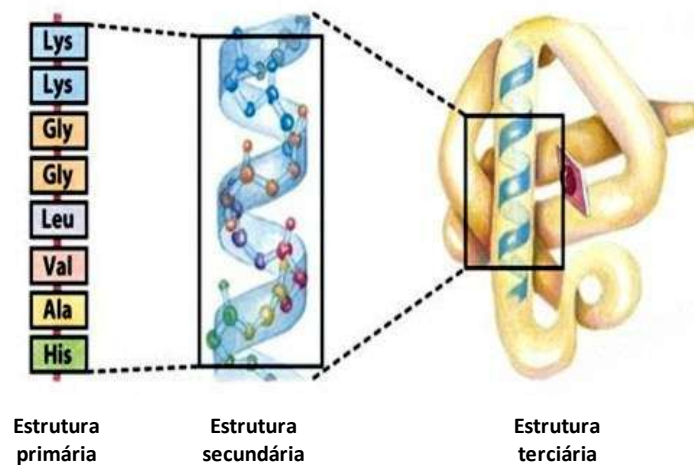
A forma de trabalho colaborativa contribuiu para que a quantidade de estruturas depositadas crescesse geometricamente. As estatísticas do PDB (Figura 1.22) demonstram o resultado do trabalho colaborativo nas quantidades de estruturas depositadas anualmente e o total acumulado de 83.768 estruturas depositadas desde a criação do RSCB PDB.



**Figura 1.22** Estatística referente às estruturas depositadas no PDB até a data de 04/09/2012, adaptado de RSCB PDB.

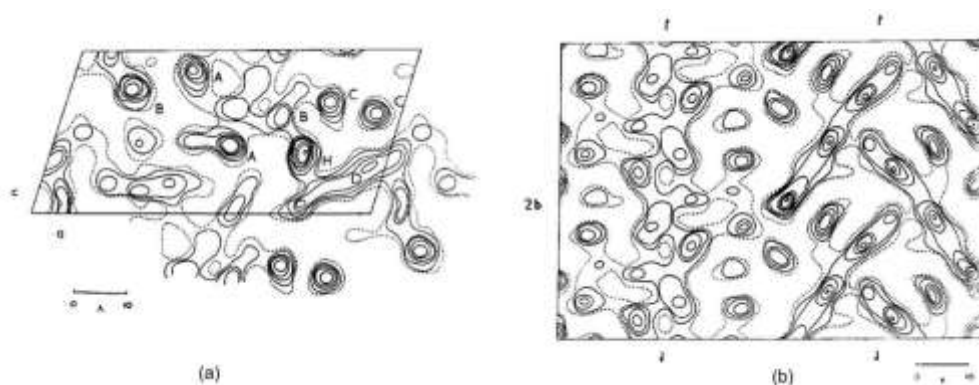
## 1.5 CIÊNCIA DA COMPUTAÇÃO E QUÍMICA SUPRAMOLECULAR

A estrutura tridimensional obtida por Kendrew e colaboradores (1958) pode ser considerada como um marco da química supramolecular, quando este concluiu que as propriedades físicas e químicas de uma proteína pudessem ser compreendidas, além das estruturas primárias e secundárias, as suas estruturas terciárias deveriam ser conhecidas uma vez que representam a macromolécula como um objeto tridimensional em uma posição no espaço (Figura 1.23).



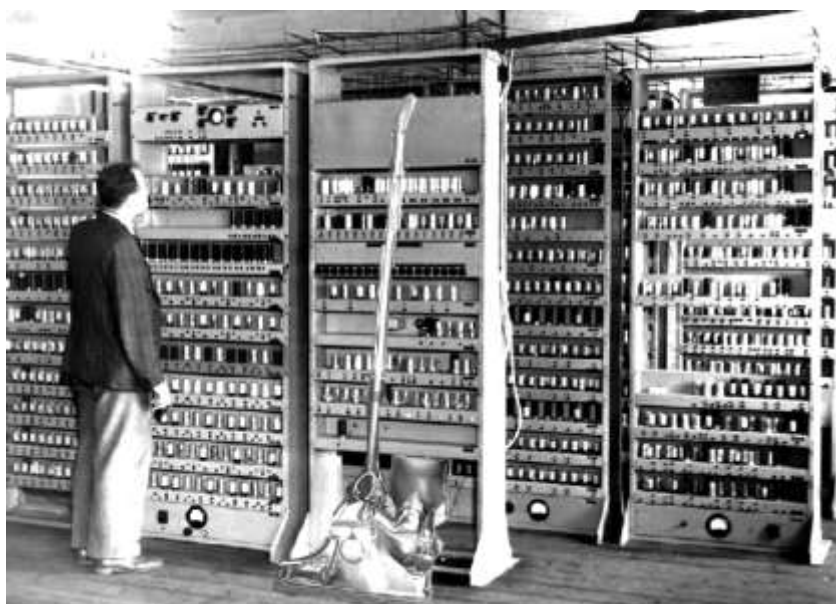
**Figura 1.23** Representação esquemática das estruturas primárias, secundárias e terciárias de uma proteína, adaptado de BIT 603 (Caracelli e Zukerman-Schpector, 2009).

Kendrew et al. (1958) utilizaram a série de Fourier para tratar matematicamente as reflexões dos raios X do ensaio cristalográfico que mapeiam as densidades de elétrons. Como uma reflexão retorna apenas dados bidimensionais, foram efetuadas 400 incidências de raios X em espaçamentos de 6 Å. Nos locais em que estão localizados átomos, há um pico de densidade de elétrons formando ilhas de densidade. A série de Fourier interpreta a função periódica pela sobreposição de harmônicas, resultante da reflexão dos raios X nesses pontos e assim permitindo o mapeamento das densidades de elétrons e estratificando-as por camadas e com isso obtendo os pontos do eixo  $y$  perpendiculares ao plano  $x$  e  $z$ , gerando a superfície da estrutura tridimensional (Figura 1.24).



**Figura 1.24 (a)** Secção tridimensional da Série de Fourier para a mioglobina tipo A. **(b)** paralelo à secção em  $x = 0$ , mostrando cadeia polipeptídica A. Extraído de Kendrew (1958).

O processamento matemático da Série de Fourier foi efetuado pelo computador Electronic Delay Storage Automatic Calculator - EDSAC (Contributors, W., 2012a) mostrada na Figura 1.25 da Universidade de Cambridge. Esse computador que operava com uma taxa de relógio de 500 kHz, incorporou em seu projeto a máquina de von Neumann e foi o primeiro computador de memória armazenada de uso prático (9,2 kb) e que continha 87 sub-rotinas para cálculos matemáticos (Ricarte, 1999). Com esta configuração, o EDSAC demorou 70 minutos para efetuar todos os cálculos necessários para que a estrutura tridimensional da mioglobina fosse obtida (Kendrew et al.; 1958). Apesar do tempo de processamento envolvido ser extremamente longo para os padrões atuais, essa foi uma experiência inovadora na utilização de processamento eletrônico para a determinação de estruturas moleculares complexas que possibilitou que a tecnologia emergente da ciência da computação interagisse sinergicamente com a física e bioquímica .



**Figura 1.25** Fotografia com vista parcial do EDSAC (Contributors, 2009).

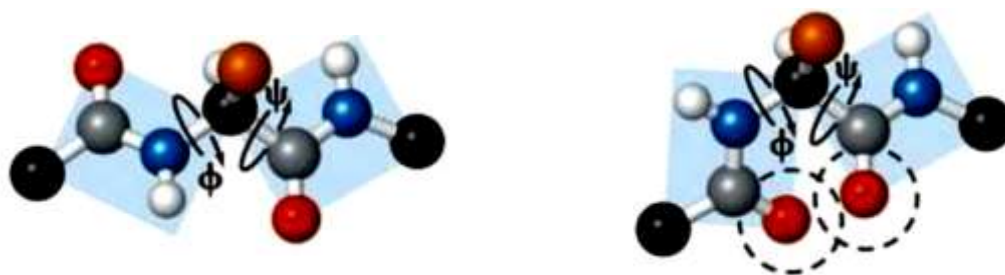
Ainda sob a premissa de que a função biológica de uma proteína está relacionada com a sua conformação estrutural, onde estão envolvidas as ligações químicas e as distâncias e ângulos envolvidos, Levinthal (Levinthal, 1966) considerou a importância da estrutura tridimensional da proteína e relacionou esta estrutura como resultado da sequência dos aminoácidos integrantes da sua cadeia polipeptídica



arranjados espacialmente para obter o menor nível de energia livre de Gibbs, Lehninger (Nelson e Cox, 2002). Dessa forma, macromoléculas com a mesma sequência de aminoácidos deveriam possuir a mesma conformação.

Considerando que os átomos presentes em uma macromolécula possuem ligações específicas com suas respectivas distâncias e os ângulos destas ligações conhecidos e com mínimas variações, observou-se a oportunidade de construção de um algoritmo que serviria de base para a criação de um programa de computador que calcularia a estrutura tridimensional da molécula.

O programa idealizado segue a sequência dos átomos em sua posição no aminoácido ao qual pertence, sendo que a cada átomo é atribuída uma coordenada espacial. Os dados de entrada do programa são os ângulos (ângulos  $\phi$  (phi) e  $\psi$  (psi)) formados entre ligações peptídicas (Figura 1.26).



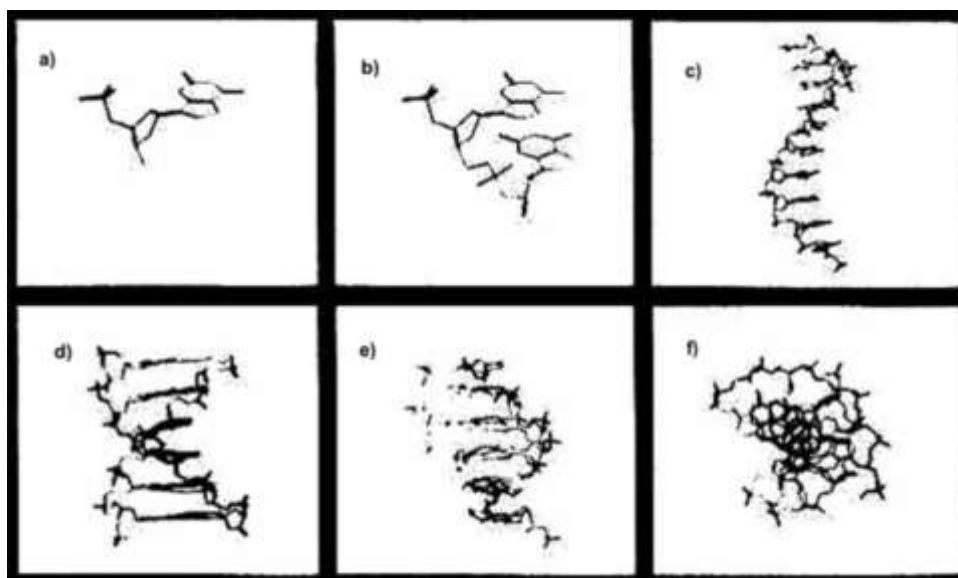
**Figura 1.26** Ângulos  $\phi$  (phi) e  $\psi$  (psi) entre ligações peptídicas, extraída de BIT 603 (Caracelli e Zukerman-Schpector, 2010).

Conforme os átomos são adicionados o programa calcula as coordenadas dos átomos em função das ligações, distâncias e ângulos envolvidos e ao final, com base nas coordenadas calculadas, é possível a construção de um modelo tridimensional de uma proteína. Todavia, observou-se que o maior problema foi a inserção dos ângulos  $\phi$  (phi) e  $\psi$  (psi) das ligações entre aminoácidos, cujos valores são variáveis e que a somatória das combinações possíveis totalizava uma quantidade extremamente elevada de saídas do programa, a maioria sem significado biológico, pois a questão de mínima energia livre de Gibbs, Lehninger (Nelson e Cox, 2002) do sistema não está contemplada no algoritmo.

Para solucionar esse problema, foi idealizada a iteração homem-máquina na qual o cientista poderia a cada passo específico do programa inserir o valor da variável desejada, ou seja, os ângulos  $\phi$  (phi) e  $\psi$  (psi) das ligações entre aminoácidos que seriam obtidas pelo cálculo em programas auxiliares. Todavia, para que ocorresse de forma eficiente, essa iteração deveria ser fácil de ser observada e assim foi criada a primeira interface homem máquina (IHM).

Foi adaptado ao computador um osciloscópio que traduzia em imagem a saída do processamento, permitindo que além de se calcular o estado de energia livre pelo programa auxiliar, também fosse visualizada e analisada de forma rápida a estrutura resultante após a inclusão das variáveis.

Com isso, estruturas de macromoléculas puderam pela primeira vez serem construídas e modeladas por computador (Figura 1.27).



**Figura 1.27** Estrutura de DNA sendo modelada pelo computador. **(a)** Inserção do nucleotídeo inicial, **(b)** Inserção do segundo nucleotídeo, **(c)** Adição de mais nucleotídeos formando uma fita de DNA, **(d)** sobreposição de duas fitas, formando uma estrutura completa de DNA, **(e)** e **(f)** Rotação da estrutura de DNA visualizando sua conformação tridimensional. Adaptado de Levinthal (1966).

O avanço da tecnologia dos computadores e softwares aplicativos trouxeram progressos na cristalografia, resultando na cristalografia computacional (Cranswick, 2008). Essa relação multidisciplinar entre áreas do conhecimento propiciou em um primeiro

momento a obtenção pela análise da difração de raios X nas coordenadas espaciais  $x$ ,  $y$  e  $z$  dos átomos de uma proteína. Isso representou um novo marco no estudo de macromoléculas, pois as coordenadas espaciais dos átomos permitem obter informações sobre a sua conformação real, que de certa forma serviram para validar os conceitos de modelagem computacional de Levinthal (1966).

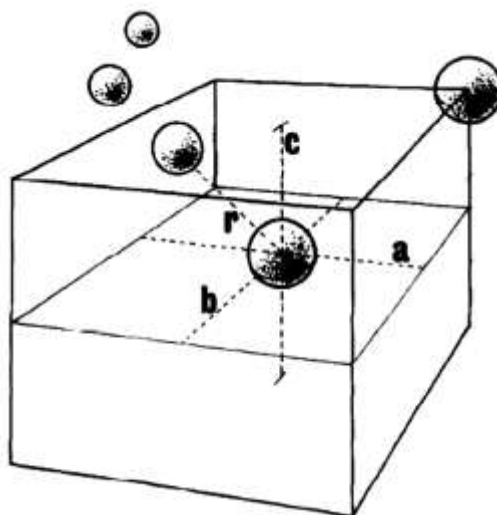
Utilizando conceitos vetoriais (Steinbruch, 1987; Leithold, 1994) a partir das coordenadas espaciais  $x$ ,  $y$  e  $z$  dos átomos de uma proteína, Tometsko (1970) idealizou um programa na linguagem FORTRAN (IBM, 1954) que pudesse mensurar as distâncias interatômicas e com isso avaliar as posições relativas de seus átomos e por meio desse artifício analisar as interações entre estes.

O programa efetua a leitura das coordenadas  $x$ ,  $y$  e  $z$  dos átomos e suas respectivas propriedades, sendo que no caso de proteínas a sequência de leitura partiria da extremidade N-terminal para C-terminal, seguindo a ordem N, CA, CB...(cadeia lateral), C, O (Figura 1.28).

IDENTIFICAÇÃO	X	Y	Z	ÁTOMO	AMINOÁCIDO
N	331.050	508.930	300.610	N.4	VAL1
CA	345.100	508.470	304.820	C.3	VAL1
C	349.580	494.280	308.210	C.2	VAL1
O	346.350	488.490	318.600	O.2	VAL1
CB	349.260	518.310	315.860	C.3	VAL1
CG1	364.480	519.030	316.820	C.3	VAL1
CG2	343.600	532.320	313.520	C.3	VAL1
H1	325.190	505.470	308.060	H	VAL1
H2	328.490	518.450	298.470	H	VAL1
H3	329.810	503.190	292.410	H	VAL1
HA	350.390	511.940	296.030	H	VAL1
HB	345.160	514.590	325.170	H	VAL1
HG11	368.540	522.390	307.360	H	VAL1
HG12	367.280	525.990	324.630	H	VAL1
HG13	368.420	509.220	319.150	H	VAL1
HG21	332.780	531.860	313.320	H	VAL1
HG22	346.790	538.880	321.520	H	VAL1
HG23	347.220	536.140	304.060	H	VAL1

**Figura 1.28** Formato de banco de dados de parte de uma proteína obtido por cristalografia, neste caso a Valina VAL1 que é o primeiro aminoácido da cadeia polipeptídica, adaptado de Karplus e Schulz (1989).

O raciocínio utilizado para determinar as interações por meio das distâncias interatômicas baseia-se no posicionamento do primeiro átomo num cubo de arestas nas dimensões  $a$ ,  $b$  e  $c$ , sendo  $a = b = c = GM$  (Figura 1.29) e  $GM$  é a variável do sistema a ser inserida pelo usuário e neste caso configurada para 3,5 Å.



**Figura 1.29** Representação esquemática do volume de controle criado pelo cubo de arestas  $a$ ,  $b$  e  $c$ , sendo  $a = b = c = GM$  que é expresso em angstroms (Å). As esferas representam os átomos da proteína. A distância  $r$  é calculada entre o átomo central e os átomos contidos no cubo. Adaptado de Tometsko (1970).

No cálculo, caso exista algum átomo posicionado dentro do volume do cubo hipotético, é calculada a sua distância  $r$  com relação ao átomo central, sendo desprezados os átomos fora do volume de controle. O programa armazena os átomos que apresentam interações e suas respectivas distâncias em formato de banco de dados e então o programa passa para o átomo seguinte na sequência convencional na molécula e assim sucessivamente aminoácido por aminoácido da cadeia polipeptídica. Uma característica do método utilizado é que pelo algoritmo utilizado, para cada átomo posicionado no centro do cubo estão envolvidas  $10^3$  operações matemáticas.

O objeto de estudo de seu trabalho foi uma molécula de mioglobina e as saídas do processamento podem ser vistas na Figura 1.30:

INUM <sup>a</sup>	BMIN <sup>b</sup>	BATOM <sup>c</sup>	DISTANCE <sup>d</sup>	CATOM <sup>c</sup>	CMIN <sup>b</sup>	JNUM <sup>a</sup>
6	GLU	CD	4,02	N	SER	3
6	GLU	CD	4,50	CA	SER	3
6	GLU	CD	4,24	CB	SER	3
6	GLU	CD	3,43	OG	SER	3
6	GLU	CD	3,30	NZ	LYS	133
<b>INUM</b>	<b>BMIN</b>	<b>BATOM</b>	<b>DISTANCE</b>	<b>CATOM</b>	<b>CMIN</b>	<b>JNUM</b>
6	GLU	OE1	4,04	C	LEU	2
6	GLU	OE1	3,85	CA	LEU	2
6	GLU	OE1	3,29	N	SER	3
6	GLU	OE1	4,09	CA	SER	3
6	GLU	OE1	3,51	CD	LYS	133
6	GLU	OE1	3,69	NZ	LYS	133
<b>INUM</b>	<b>BMIN</b>	<b>BATOM</b>	<b>DISTANCE</b>	<b>CATOM</b>	<b>CMIN</b>	<b>JNUM</b>
e → 6	GLU	OE2	3,80	CD	LYS	133
<b>6</b>	<b>GLU</b>	<b>OE2</b>	<b>3,55</b>	<b>NZ</b>	<b>LYS</b>	<b>133</b>
<b>INUM</b>	<b>BMIN</b>	<b>BATOM</b>	<b>DISTANCE</b>	<b>CATOM</b>	<b>CMIN</b>	<b>JNUM</b>
7	TRY	O	4,47	O	LEU	9
7	TRY	O	3,58	C	LEU	9
7	TRY	O	3,08	N	LEU	9
7	TRY	C	3,47	C	VAL	10
7	TRY	C	2,91	N	VAL	10
7	TRY	O	3,37	CA	VAL	10
→ <b>7</b>	<b>TRY</b>	<b>C</b>	<b>2,64</b>	<b>N</b>	<b>LEU</b>	<b>11</b>

**Figura 1.30** Output do programa do computador pela análise sequencial dos átomos de uma molécula de mioglobina, extraído de Tometsko (1970).

a – INUM e JNUM indicam a posição do aminoácido na cadeia polipeptídica.

b – BMIN e CMIM indicam os aminoácidos que respectivamente ocupam as posições INUM e JNUM.

c – BATOM e CATOM são as denominações dos átomos dos aminoácidos, sendo que BATOM especifica os átomos posicionados no centro do volume de controle.

d – Expressa as distâncias calculadas entre o átomo central ao cubo e outro posicionado internamente ao volume de controle (INUM e JNUM).

As setas indicam as menores distâncias calculadas para os átomos centrais.

Os cálculos efetuados permitem a identificação das interações entre os átomos de diferentes aminoácidos pertencentes à cadeia polipeptídica e assim identificar as regiões onde ocorrem os enovelamentos. A localização é possível pela depuração dos dados obtidos na Figura 1.30 de forma a gerar um relatório analítico com as interações consideradas válidas. Com base nos átomos envolvidos e suas respectivas distâncias torna-se possível a construção de um modelo tridimensional.

A evolução desse método com relação ao proposto por Kendrew deve-se ao fato da localização dos átomos em coordenadas espaciais  $x$ ,  $y$  e  $z$  resultantes dos ensaios cristalográficos passarem a ser disponibilizados de forma padronizada com o auxílio de sistemas computacionais.

Importante ressaltar que a quantidade de operações matemáticas envolvidas nesse método é diretamente proporcional ao valor de GM (aresta do cubo). Ao incrementar o seu valor de 3,5 Å para 4,0 Å, o volume de dados aumentou de 23.000 linhas para 41.000 linhas. Assim, como estabelecer o parâmetro ideal para GM torna-se dependente de análises independentes em função da proteína estudada, uma vez que proteínas diferentes apresentam enovelamentos específicos e as distâncias envolvidas nas interações entre as cadeias laterais não seguem a um padrão específico.

Ao final da década de 1960 e início da de 1970, houve o advento dos discos rígidos e o processamento de dados tornou-se mais dinâmico em função da possibilidade de acesso rápido aos dados armazenados em qualquer posição do disco rígido, o que não ocorria com as fitas magnéticas, onde o acesso aos dados era efetuado de forma sequencial ao seu comprimento. Assim, com os discos tornou-se possível a estruturação hierárquica de bancos de dados em listas e árvores binárias (Silberschatz et al., 2006a) e permitindo que estes pudessem ser manipulados com facilidade.

Uma das formas de bancos de dados são os “*arrays*” ou vetores unidimensionais, que permitem a ordenação dos dados e a execução das operações de extração e armazenamento de dados em função de um índice associado ao vetor (Lipschutz e Lipson, 2004; Tenenbaum et al., 1995). A Figura 1.31 ilustra um “*array*” ou vetor unidimensional genérico.

Índice	Dados
1	
2	
n	

**Figura 1.31** Representação de um banco de dados unidimensional, adaptado de Lipschutz e Lipson (2004).

A tecnologia disponível permitiu que Katz e Levinthal (1972) utilizassem esses recursos para modelar macromoléculas com programas e visualizassem o resultado dessa modelagem por meio de interações gráficas em computadores.

Para a modelagem da estrutura molecular faz-se necessário o cálculo das coordenadas espaciais dos átomos na molécula considerando suas possíveis interações. Assumiu-se que os átomos estão ligados aos seus vizinhos por meio de ligações químicas e essas ligações com distâncias fixas e o resultado desse arranjo apresentando o menor nível de energia livre de Gibbs (Nelson e Cox, 2002) e podendo ser posteriormente o modelo comparado a estruturas obtidas experimentalmente pelo uso de difração de raios X ou outros meios. O conceito de modelagem molecular por computador com base nas ligações químicas entre os átomos e seus vizinhos foi anteriormente explorado por Levinthal e Tometsko (Levinthal, 1966; Tometsko, 1970), sendo que neste trabalho Katz e Levinthal (1972) utilizaram conceitos de bancos de dados (Lipschutz e Lipson, 2004; Silberschatz et al., 2006a; Tenenbaum et al., 1995) e matemática discreta (Lipschutz e Lipson, 2004) para obter o CHEMGRAPH na linguagem FORTRAN (IBM, 1954), que é composto de várias sub-rotinas, sendo que as 11 principais estão listadas a seguir:

1.	CMGRF	Programa principal que interpreta os comandos e aciona as sub-rotinas.
2.	MOLE	Aciona a função de construção molecular, conectividade, cálculos de coordenadas e estrutura da árvore (Lipschutz e Lipson, 2004).
3.	JIGGLE	Manipulação interna da molécula para translações e rotações.
4.	FILING	Rotinas de gerenciamento de arquivos.
5.	SOLVE	Cálculo para adequação de mínima energia e mapa de energia.
6.	WIRE	Formação da imagem e output para visualização.
7.	LINKIO	Software de alta velocidade para o visualizador ADAGE AGT/50.
8.	PDSPL	Interpretador de comandos e sistema de imagem.
9.	CUBE	Posicionamento dos átomos na estrutura de acordo com sua localização espacial.
10.	CHMAN	Rotina para manuseio de caracteres e parâmetros.
11	WISWR	Processador para a notação química linear Wiswesser (Fritts e Schwind, 1982).

O programa foi modelado para trabalhar com bancos de dados estruturados em “*arrays*” unidimensionais (Lipschutz e Lipson, 2004; Tenenbaum et al., 1995) que contém as informações sobre as moléculas, os átomos que as formam ou fragmentos moleculares. Durante o processamento, ao serem formadas moléculas ou fragmentos destas, os resultados são armazenados em um banco de dados de forma que o programa

possa utilizá-los na composição de estruturas complexas. A estruturação desse primeiro banco de dados pode ser vista na Figura 1.32.

1. TYPE	The type number of the atoms. This type number is the index of the atom type in the ATYPES arrays.
2. VIRT	If $VIRT(I)=0$ , $I$ is a real atom. If $VIRT(I)=J>0$ , $I$ is a ring closure virtual atom. Its type and coordinates should be equal to that of atom $J$ . If $VIRT(I)=K<0$ , $I$ is a vanishing virtual representing the $-K$ th unsatisfied bond in the molecule.
3. REF	The absolute value of $REF(I)$ indicates which subunit atom $I$ is in. $REF(I)$ indexes the array of subunits, UNITNM. If $REF(I)$ is negative, the atom is a reference atom.
4. FROM	$FROM(I)$ is the atom from which atom $I$ comes. The constraint that each atom follows no more than one other implies a tree structure on the graph of the molecule. Where rings are wanted, ring closure virtuals must be employed. If $FROM(I)=0$ , atom $I$ begins a new disjoint segment.
5. X	$X(I)$ , $Y(I)$ , $Z(I)$ are the coordinates in angstroms for atom $I$ .
6. Y	
7. Z	
8. THETA	The rotation angle of the bond between atom $I$ and atom $FROM(I)$ .
9. BOND	$BOND(I)$ is the valence of the bond between $I$ and $FROM(I)$ . If $BOND(I)$ is negative, the bond is considered nonrotatable (locked against rotation). Bonds of valence 2 or greater are normally locked. Molecular conformation is changed by rotation about unlocked bonds.
10. LEVEL	A level indicator. Each atom in the main chain has level 1. Side chains branching directly off the main chain have level 2 and so on. Note that if $LEVEL(I)$ is equal to $LEVEL(FROM(I))$ , then $I$ is the highest numbered atom coming from $FROM(I)$ .
11. ANGLO	$ANGLO(I)$ is the highest numbered unlocked bond whose rotation affects the position of atom $I$ . Note that if two atoms have equal ANGLOS, the distance between them is fixed. ANGLO must be recomputed each time the BOND array is altered.
12. QBACK	$QBACK(I)$ is the number of the atom back down the chain such that closer interactions will not contribute to the conformation energy on a pair-wise basis. This variable will be used only for macromolecules.

Figura 1.32 Estrutura de banco de dados para moléculas, extraído de Katz e Levinthal (1972).

Para que o programa pudesse manipular o banco de dados da figura acima, foi necessário que este fosse estruturado para trabalhar com outros bancos de dados, que possuem a função auxiliar de armazenando informações referentes aos átomos constituintes das moléculas e suas características com relação à valência das ligações e seus ângulos, reproduzidas nas figuras abaixo (Figuras 1.33 e 1.34).


















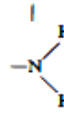







Number	Symbol	Description	Chemical representation	Bond class
1	VV	Single-bonded vanishing virtual	—VV	1
2	DVV	Double-bonded vanishing virtual	=DVV	1
3	TVV	Triple-bonded vanishing virtual	≡TVV	1
4	H	Hydrogen	—H	1
5	CL	Chlorine	—CL	1
6	C	Carbon, tetrahedral		4
7	RC	Carbon, resonant (as in benzene)		3
8	DC	Carbon, double-bonded		3
9	TC	Carbon, triple-bonded		5
10	O	Oxygen, two single bonds		2
11	DO	Oxygen, double-bonded		1
12	N	Nitrogen, three single bonds		3
13	DN	Nitrogen, double-bonded		2
14	TN	Nitrogen, triple-bonded		1
15	P	Phosphorus, tetrahedral		4
16	S	Sulfur, analogous to 'O'		2
17	F	Fluorine	—F	1
18	BR	Bromine	—BR	1
19	I	Iodine	—I	1

Figura 1.33 Banco de dados com os átomos padrão para a construção de moléculas e suas respectivas classificações de ligações, extraído de Katz e Levinthal (1972).

Bond class	Characteristic bonding angles ( $\phi$ )	Angular difference between bonds ( $\Delta\theta$ )
1. Univalent elements and additions	—	—
2. O, S, and bivalent additions	105°	—
3. N and trivalent additions	120°	180°
Double-bonded C (DC) and heterocyclic carbon	—	—
4. C, P, and tetrahedral additions	109.5°	120°
5. Triple-bonded C (as in CC and CN) and triple bonded N (as in CN)	180°	—

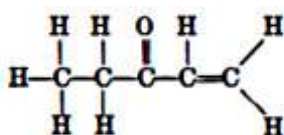
Figura 1.34 Banco de dados com os ângulos das ligações para os átomos, extraído de Katz e Levinthal (1972).

Durante a modelagem, o programa necessita a inserção da fórmula química da molécula. Para grupos comuns, há um banco de dados pré-definido com a equivalência da notação química e a notação utilizada pelo CHEMGRAPH (Figura 1.35).

Group name	Chemical notation	Equivalent CHEMGRAF notation*
CH3		C(H, H, H)
CH2		C(H, H)
CH		b
PO		b
NH2		N(H, H)
NH		N(H)
CN		TQ(TN)
CC		TC·TC
CO		DC(DO)
DCH2		DC(H, H)
DCH		DC(H)
OH		O(H)

**Figura 1.35** Banco de dados de grupos pré-definidos para a modelagem molecular, extraído de Katz e Levinthal (1972).

Uma vez que os parâmetros referentes aos átomos, suas valências e ângulos de ligações são armazenados nos diversos bancos de dados, o programa permite a construção de moléculas pela digitação da sequência de átomos e grupos pré-definidos (Figura 1.36).

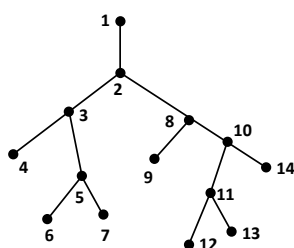


**Figura 1.36** Representação química da sequência CH<sub>3</sub>.CH<sub>2</sub>.CO.CH.C<sub>2</sub>\* onde o asterisco indica ao programa o fim da cadeia, extraído de Katz e Levinthal (1972).

O conceito de grupos pré-definidos em banco de dados é entendido para moléculas peptídicas, permitindo a modelagem de uma proteína ou parte desta. Em todos os casos, a valência do átomo responsável pela ligação deverá ser igual ao seu correspondente virtual na estrutura.

Todas as sub-rotinas do programa assumem que as ligações químicas possuem comprimentos fixos e que os ângulos entre as diferentes ligações ao mesmo átomo são também fixas. Exceção a esta regra são os anéis os quais não são calculados em função de distâncias e ângulos envolvidos e suas estruturas depositadas em uma biblioteca específica.

Quando uma sequência de átomos é formada para a modelagem de uma molécula, a quantidade de ligações, suas respectivas distâncias e ângulos determinam a sua estrutura espacial. Quando algum átomo é alterado, há a possibilidade de que a estrutura espacial da molécula modelada seja alterada. O programa utilizou o recurso de banco de dados disposto em árvores binárias (Lipschutz e Lipson, 2004; Silberschatz et al., 2006b; Tenenbaum et al., 1995) para representação da molécula (Figura 1.37), com o objetivo de que qualquer alteração efetuada em um ponto específico durante a modelagem possa corrigir a estrutura seguinte dependente das ligações e ângulos do átomo alterado.



Átomo	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Átomos Dependentes	2	3 8	4 5	-	6 7	-	-	9 10	-	11 14	12 13	-	-	-

**Figura 1.37** Representação de uma molécula no formato de árvore binária (Lipschutz e Lipson, 2004; Silberschatz et al., 2006b; Tenenbaum et al., 1995), adaptado de Katz e Levinthal (1972).

Este trabalho de Katz e Levinthal (1972) utilizou conceitos de matemática discreta aplicada à busca em bancos de dados. Por outro lado, o sistema criado para a modelagem tem como output modelos rígidos em função dos bancos de dados referentes às ligações químicas trabalharem distâncias e ângulos fixos.

## **1.6 VIZUALIZADORES GRÁFICOS**

Desde que Kendrew (1958) teorizou que para a melhor compreensão de um sistema biológico é necessário a visualização da sua estrutura, diversos autores desenvolveram suas pesquisas no intuito de obter meios de utilizar os recursos computacionais que além do avanço técnico tornaram-se mais acessíveis (Howard e Kollman, 1988) para poder processar modelos teóricos e analisar em tela gráfica não apenas a imagem das estruturas biomoleculares, mas também os tipos de interação entre seus átomos, potenciais de carga eletrostática e regiões onde estão localizados os sítios ativos (Ferrin et al., 1988), dentre outras.

Atualmente a utilização de softwares para visualização de macromoléculas em computadores pessoais é bastante comum (Herráez, 2006), haja vista a popularização deste equipamento com relação aos custos de aquisição e a constante capacidade de processamento (Howard e Kollman, 1988). A visualização gráfica permite uma melhor visualização de imagens de propriedades que anteriormente só eram analisadas numericamente. Segundo Voet e Hammons (2005), a utilização de interfaces gráficas é vital para o estudo das proteínas e suas estruturas.

Algumas vantagens dos softwares para visualização gráfica de macromoléculas, de acordo com Herráez (2006) são:

- a- Permitem a utilização de diversas representações de uma molécula, como wireframe, sticks, ball and sticks, CPK (preenchimento de espaço) e para as proteínas e DNA a representação esquemática renderizada.
- b- A percepção tridimensional da estrutura, obtida por sua rotação em qualquer dos seus eixos, bem como a noção de perspectiva e profundidade.

c- A interação com o usuário, onde este pode escolher as ações que devem ser executadas, tais como rotação, seleção e configuração do tipo de representação de parte da estrutura, medições diversas envolvendo as propriedades geométricas da estrutura, suas moléculas e interações destas e dos seus átomos.

O primeiro desses softwares desenvolvido e que se tornou popular foi o RasMol (Massachusetts, 1995; Martz e Francoeur, 1997; Herráez, 2006) que é a junção de **Raster** (a matriz de pixels em uma tela de computador) com **Molecule** (Figura 1.38). O programa gera uma imagem de computador de uma molécula com base nos dados das coordenadas existentes nos arquivos PDB (RSCB; EMBL-EBI, 1995; Sussman et al., 1999; Berman et al., 2000; Voet e Hammons, 2005; Berman et al., 2007; Sussman et al., 2012) que são processados para obter uma combinação de profundidade e sombreamento que permite que a imagem na tela assuma propriedades tridimensionais.

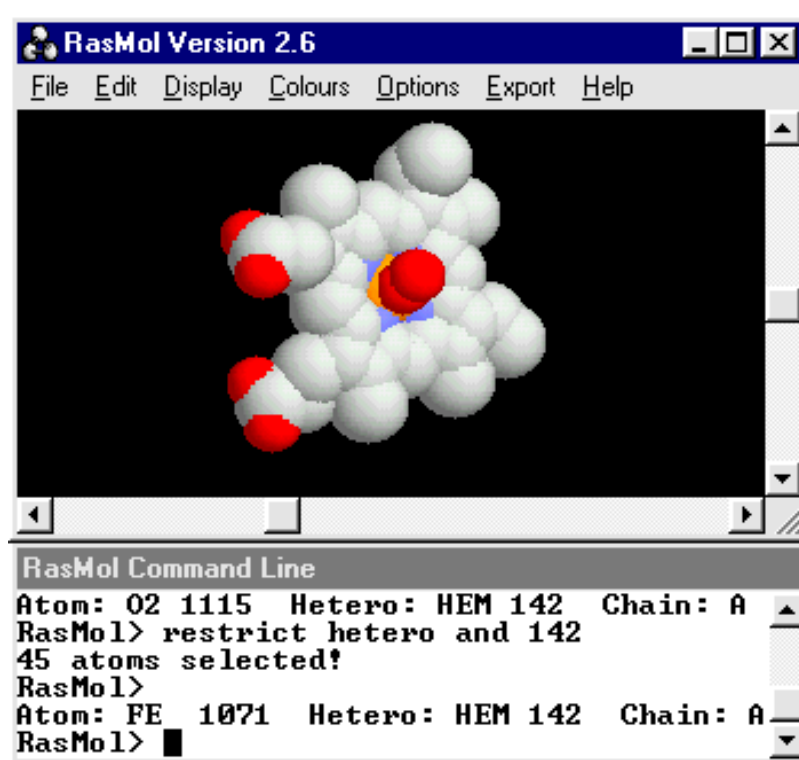
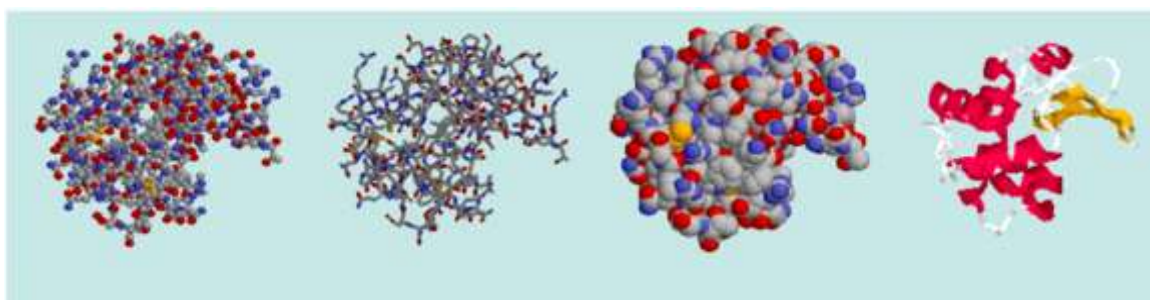


Figura 1.38 Imagem capturada de uma tela com o programa RasMol em uso, extraído de Martz (1996).

Logo após, em 1996 foi criado por Bryan van Vliet e Tim Maffett o Chime (**CH**emical **mIME**) (Martz, 1996; 2004; Herráez, 2006; Martz, 2012a; b) que é um aprimoramento do RasMol (Massachusetts, 1995; Martz, 1996; Martz e Francoeur, 1997; Herráez, 2006), sendo que enquanto o RasMol é processado no próprio computador e utiliza as coordenadas atômicas das estruturas depositadas no PDB (RSCB; EMBL-EBI, 1995; Sussman et al., 1999; Berman et al., 2000; Voet e Hammons, 2005; Berman et al., 2000; 2007; Bioinformatics, 2012; Sussman et al., 2012), o Chime (Figura 1.39) é um visualizador gráfico que opera como um aplicativo dentro do navegador da Internet e seu uso somente é possível se o autor de uma estrutura molecular a deposita em uma página que possa ser acessada. Atualmente este programa está sendo descontinuado devido ao fato de que os seus criadores recusaram-se a abrir o seu código-fonte, impedindo uma maior participação da comunidade científica no seu desenvolvimento.



**Figura 1.39** Imagens da lisozima 1HEL capturada de uma tela com o programa Chime em uso, extraído de Martz (2012b).

Em 2004 surgiu o Jmol (Herráez, 2006; Contributors, 2012; Jmol, 2012) como sucessor do Chime (Martz, 1996; 2004; Herráez, 2006; Martz, 2012a; b), sendo este um projeto open source que aproveitou em grande parte as linguagens de comando do Chime/Rasmol. Este software possui como característica operacional a plataforma Java e não necessita de instalação, podendo ser processado em navegadores dos sistemas operacionais como o Mac OSX (Safari) ou Linux (Firefox) como faz no MS Windows (Internet Explorer 7).

A versatilidade deste software pode ser observada na quantidade de formatos arquivos de bancos de dados de proteínas que consegue processar (48 formatos no total) e estando os mais conhecidos relacionados abaixo (Tabela 1.5):

**Tabela 1.5** Formatos de arquivos processados pelo Jmol, adaptado de Jmol (2012)

MOL	Estrutura MDL / Elsevier / Symyx (versão clássica V2000)
PDB	Protein Data Bank - Research Collaboratory for Structural Bioinformatics
XYZ	arquivo XYZ format, XMol - Minnesota Supercomputer Institute
MOL2	Sybyl, Tripos
GAMESS	General ATomic and Molecular Electronic Structure System output (tanto para as variantes USA e Reino Unido) - Gordon Research Group, Iowa State University
SMOL	Spartan data - Wavefunction, Inc.
PQR	Formato PDB modificado, incluindo carga e raios atômicos.
Amber	The Amber, suite de programas de simulação de modelagem molecular.
XSD	Accelrys Materials Studio
AMPAC	AMPAC output - Semichem, Inc.

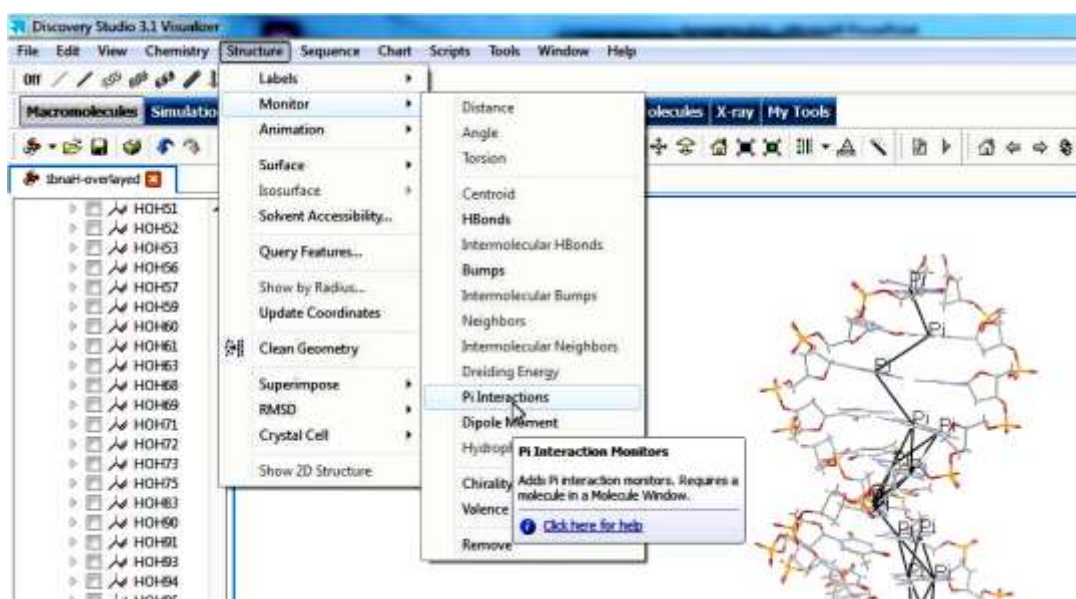
Este software para visualização gráfica está incorporado nas páginas do PDBsum e do PDB-RSCB (Embl-Ebi, 1995; RSCB), permitindo que por meio da internet o usuário possa analisar por meios gráficos a estrutura tridimensional de uma proteína (Figura 1.40).

**Figura 1.40** Visualização on-line da estrutura tridimensional da proteína glutathione reductase 1GRE utilizando o Jmol, extraído de PDB (RSCB).

Após esse longo caminho evolutivo no qual tem-se perseguido a meta da interação entre o pesquisador e a macromolécula por meio de interfaces gráficas, onde o ambiente virtual permite uma análise das propriedades de cada estrutura em particular, não tardou que os softwares de visualização saíssem dos centros de pesquisa das universidades para a iniciativa privada.

Existe em uso pelo LaCrEMM e BioMat a versão freeware do software desenvolvido pela empresa Accelrys Software Inc.. Trata-se do Discovery Studio Client 3.1<sup>TM</sup> (Accelrys, 2011) que além de processar modelos teóricos e analisar em tela gráfica não apenas a imagem das estruturas biomoleculares analisa também os tipos de interação entre seus átomos, potenciais de carga eletrostática, regiões dos sítios ativos e ligações de hidrogênio e inovou nesta versão com a disponibilização do cálculo de interações  $\pi$ .

O algoritmo utilizado permite a identificação das interações  $\pi$  $\pi$ ,  $\text{cátion}\pi$  e  $\text{sigma}\pi$ , e conseqüentemente a construção das suas representações gráficas em tela (Figura 1.41) de forma a permitir uma análise com maior detalhamento das interações existentes nas macromoléculas.



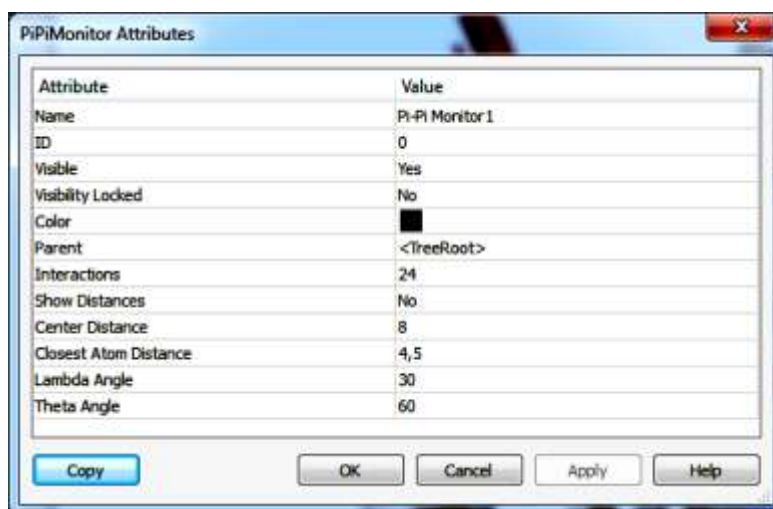
**Figura 1.41** Visualização da estrutura tridimensional do DNA 1BNA com as interações  $\pi$  $\pi$ , utilizando o software Discovery Studio Client 3.1<sup>TM</sup> (Accelrys, 2011).



Para o cálculo dos tipos de interação  $\pi$  o algoritmo localiza todos os anéis onde estas ocorrem. Os ácidos aminados fenilalanina, histidina, tirosina e triptofano são localizados pela sua estrutura. Para outros resíduos e pequenas moléculas, um anel que pode realizar uma interação  $\pi$  é definido como sendo um anel plano com hibridização  $sp^2$ . Isto inclui anéis aromáticos, mas também não aromáticos a exemplo das bases do DNA.

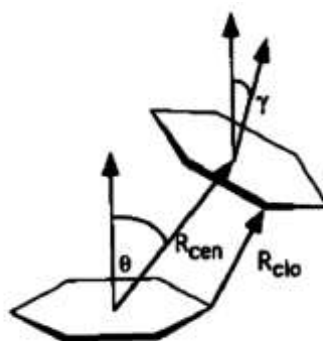
Embora o software possua atributos pré-definidos para calcular a ocorrência de interações  $\pi$ , é possível a alteração desses parâmetros por meio de uma caixa de diálogo específica.

Os parâmetros envolvidos nos cálculos e que podem ser alterados pelo usuário em função da especificidade da sua pesquisa estão contidos em uma caixa de diálogo conforme a Figura 1.42 a seguir:



**Figura 1.42** Caixa de diálogo para alteração dos atributos envolvidos no cálculo das interações  $\pi$ , extraído de Discovery Studio Client 3.1<sup>TM</sup> (Accelrys, 2011).

Para o cálculo das interações  $\pi$  a Accelrys Software Inc. adotou a metodologia de McGaughey (Mcgaughey et al., 1998), mostrado na Figura 1.43.



**Figura 1.43** Empilhamento de dois anéis envolvidos em uma interação  $\pi\cdots\pi$ , onde  $R_{cen}$  é a distância entre os dois centros;  $R_{clo}$  é a menor distância existente entre os átomos dos dois anéis,  $\theta$  o ângulo entre o vetor normal ao plano do anel e o vetor distância entre os centros e por fim  $\gamma$  o ângulo formado pelos vetores normais aos planos de cada anel, extraído de McGaughey (Mcgaughey et al., 1998).

Embora os parâmetros da caixa de diálogo para o cálculo das interações  $\pi$  do Discovery Studio Client 3.1 possam ser alterados, em função da metodologia definida por McGaughey e colaboradores McGaughey (Mcgaughey et al., 1998), o software considera em seus cálculos os seguintes padrões:

- Distância  $R_{cen}$  : 8,0 Å máximos
- Distância  $R_{clo}$  : 4,5 Å máximos
- Ângulo  $\theta$  : entre  $0^\circ$  e  $60^\circ$
- Ângulo  $\gamma$  : entre  $0^\circ$  e  $30^\circ$

Para as interações  $\text{cátion}\cdots\pi$ , o software efetua uma busca para encontrar cátions que com relação ao centroide atendam aos parâmetros abaixo:

- Distância  $R_L$  : 6,0 Å máximos
- Ângulo  $\theta$  :  $< 45^\circ$

Sendo  $R_L$  a distância entre o centroide e o cátion e  $\theta$  o ângulo formado entre o vetor normal ao plano do anel e o vetor distância da interação  $\text{cátion}\cdots\pi$ .

Para as interações  $\sigma\cdots\pi$ , o software efetua uma busca para encontrar átomos de carbono que estejam ligados a átomos de hidrogênio de forma explícita ou não explícita e que com relação ao centroide atendam aos parâmetros abaixo:

- Distância  $R_L$ : 4,0 Å máximos
- Ângulo CH-centroide :  $< 20^\circ$
- Ângulo vetor CH e o plano do anel:  $< 45^\circ$

Considera-se na busca dos átomos de hidrogênio que estão envolvidos nas interações  $\sigma\cdots\pi$ , aqueles que satisfaçam as condições acima e que estejam a uma distância 0,5 Å menor que a distância  $R_L$ , sendo  $R_L$  a distância relativa ao centroide.

## 1.7 CONCLUSÕES DO CAPÍTULO 1

Neste capítulo, além da caracterização dos parâmetros geométricos para os diversos tipos de interações  $\pi$ , foi efetuada uma retrospectiva histórica dos avanços da química supramolecular e bioquímica onde a ciência da computação contribuiu sinergicamente para os resultados obtidos. A criação do PDB, com seu banco de dados padronizado para as estruturas tridimensionais, favoreceu a construção de visualizadores gráficos que permitem o estudo das biomoléculas e algumas de suas propriedades em telas de computadores. Todavia, a identificação das interações  $\pi$  não é efetuada de forma exata nos softwares existentes, o que é um problema a ser solucionado. Como solução para este problema, idealizou-se que uma vez conhecidos os parâmetros geométricos para as interações  $\pi$  é possível identificá-las em qualquer estrutura de biomoléculas depositada no PDB utilizando cálculos de álgebra linear processados em um computador. Esse conceito é abordado no capítulo a seguir.

# Capítulo 2

---

MATERIAIS E MÉTODOS

O objetivo deste capítulo é apresentar os métodos utilizados nos quais os conceitos matemáticos da álgebra vetorial são associados às coordenadas das estruturas moleculares obtidas por ensaios cristalográficos depositadas no PDB. Um ponto central do trabalho é o uso dos cálculos vetoriais processados em computador para a determinação das coordenadas de centroides, obtenção do vetor normal ao plano do anel aromático, o vetor distância e seu módulo entre o centroide e demais átomos ou outros centroides da estrutura e o ângulo resultante entre estes e o vetor normal ao plano do anel aromático. Os procedimentos descritos neste capítulo são a solução em resposta ao problema da falta de uma ferramenta que determine com exatidão as interações  $\pi$ , o que vem a constituir uma inovação para o estudo dessas interações em biomoléculas. Para ilustrar as aplicações destes conceitos foram construídas figuras dos aminoácidos Fenilalanina (Phe), Histidina (His), Triptofano (Trp) e Tyrosina (Tyr), com as bases do DNA<sup>1</sup> Adenina (A), Citosina (C), Glutamina (G) e Timina (T). Ao final do capítulo há uma descrição de como utilizar planilhas eletrônicas para a realização dos cálculos envolvidos, selecionar os resultados em função dos parâmetros considerados (comprimento e ângulo para cada tipo de interação  $\pi$  e como exportar o arquivo resultante desse processo em um formato que possa ser utilizado por um visualizador gráfico.

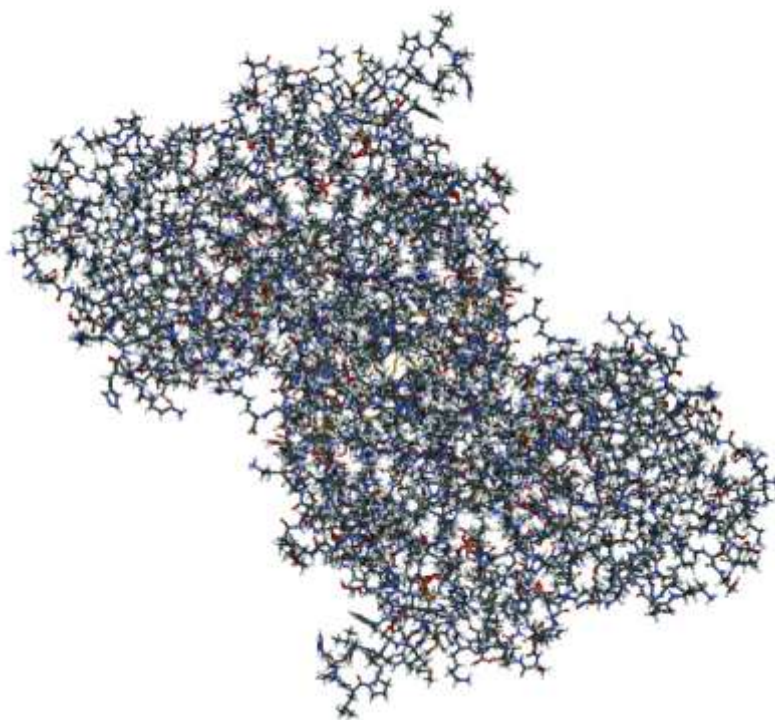
## 2.1 ÁLGEBRA VETORIAL E QUÍMICA SUPRAMOLECULAR

Segundo Menon (2009), o conceito vetorial possui relevante importância nas ciências exatas e está associado a grandezas que necessitam de módulo, direção e sentido para serem completamente especificadas. A álgebra linear, empregando os conceitos vetoriais permite equacionar problemas trigonométricos associados a essas grandezas, e no caso do estudo das macromoléculas pode ser utilizado como base para um fluxograma para mapear as interações  $\pi$  em função de variáveis de contorno ou características

---

<sup>1</sup> As siglas das bases do DNA no PDB apresentam o prefixo “D”. Este prefixo serve para indicar que esta base pertence a uma molécula de DNA e não de RNA. Em função dessa convenção nas estruturas estudadas é adotada a seguinte terminologia: Adenina (DA), Citosina (DC), Glutamina (DG) e Timina (DT).

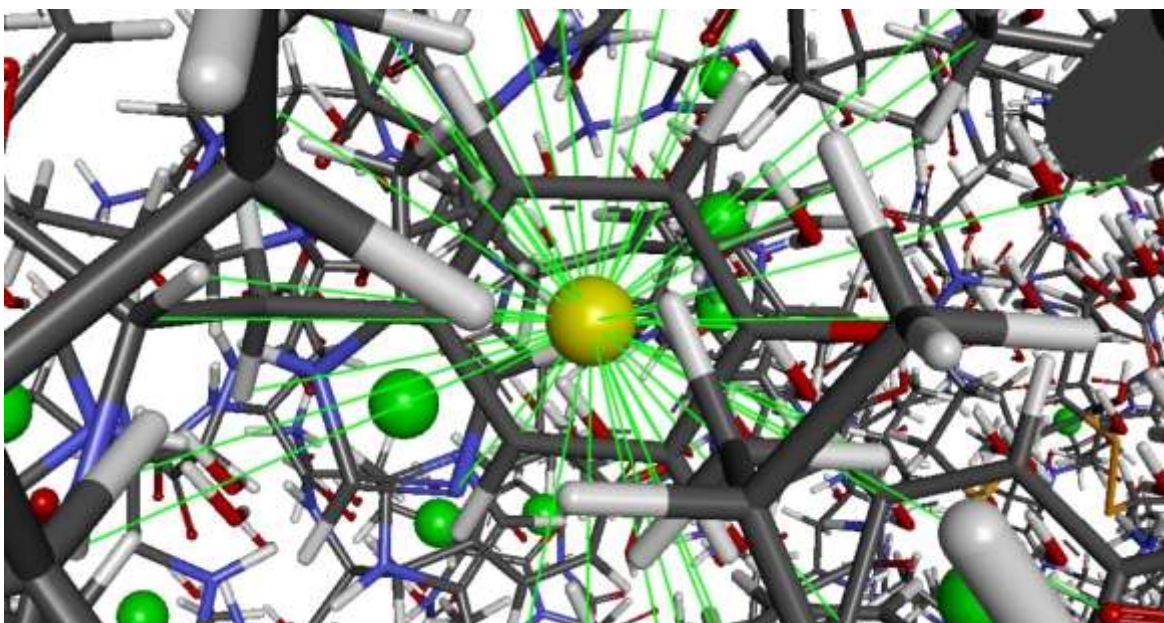
especiais relacionadas ao comprimento das interações e o seu ângulo relativo ao vetor normal ao plano conforme especificados na Figura 1.9, (Zukerman-Schpector e Haiduc, 2002; Tiekink e Zukerman-Schpector, 2009; Tiekink e Zukerman-Schpector, 2010; Zukerman-Schpector et al., 2011; Caracelli, et al., 2010; Caracelli, Zukerman-Schpector, Tiekink, 2012; Tomić et al., 2006; Babu, 2003). Dada a pequena quantidade de moléculas com anéis aromáticos e com hibridização  $sp^2$  que podem realizar uma interação  $\pi$ , dentre eles os aminoácidos fenilalanina, histidina, tirosina e triptofano, as bases do DNA e ligantes que satisfaçam essas condições (Seng et al., 2010; Ermondi e Caron, 2006; Mignon et al., 2005; Tewari e Dubey, 2008; McGaughey et al., 1998), um método matemático baseado em cálculo vetorial viria a permitir a identificação de forma automática dessas interações, principalmente em moléculas de proteínas como a da Figura 2.1, que possui em torno de 14.500 átomos e que contem apenas 110 anéis entre aminoácidos e ligantes onde podem ocorrer interações  $\pi$ .



**Figura 2.1** Representação gráfica de uma glutathione redutase código PDB 1GRA com seus átomos. Estrutura gerada a partir de dados cristalográficos (Karplus e Schulz 1989), visualizada com a utilização do software Discovery Studio 3.1<sup>TM</sup>.

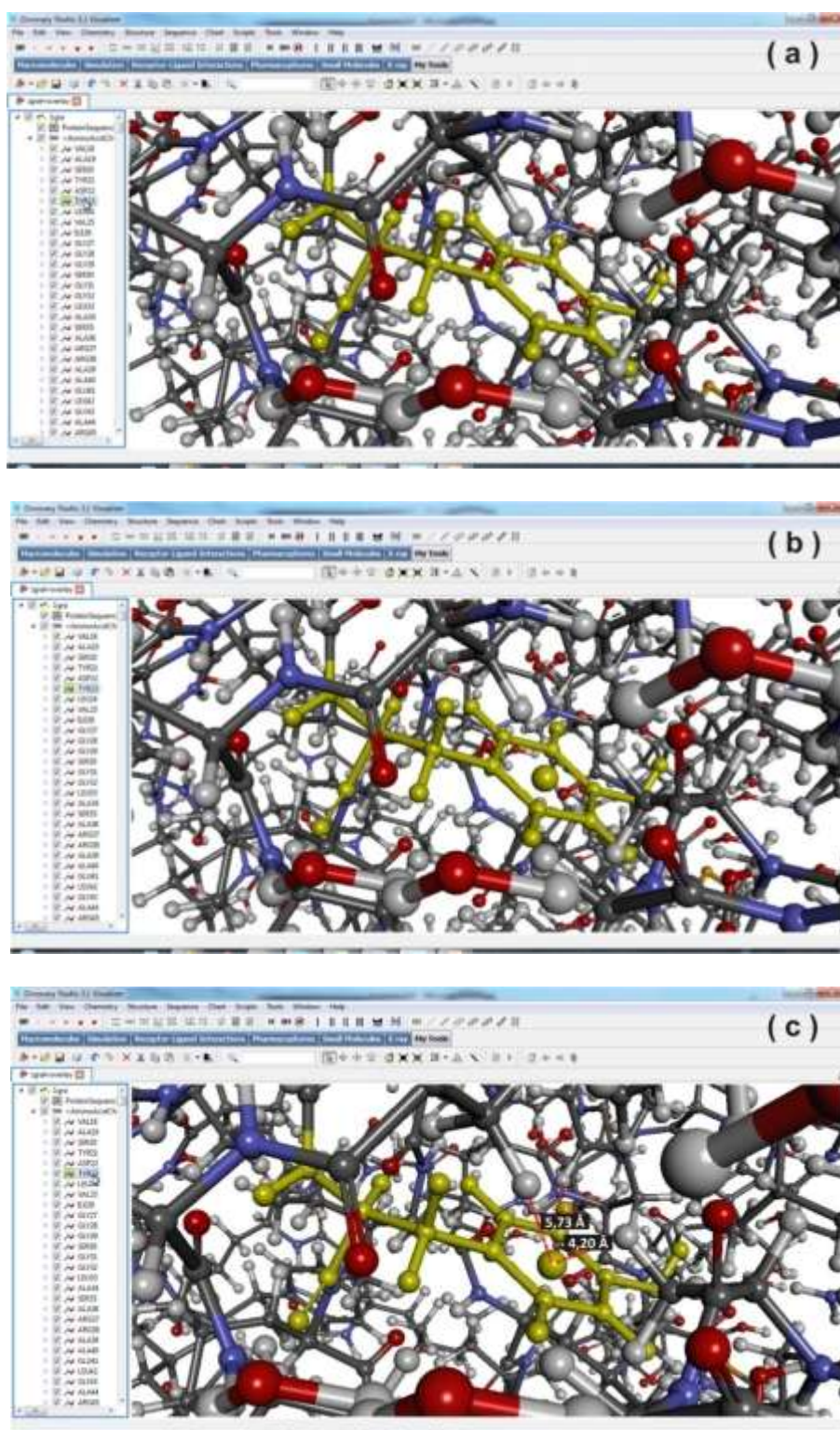
Considerando 4,5 Å como sendo o comprimento máximo de uma interação  $\pi$ , verifica-se que os átomos estão localizados a esta distância máxima de um centroide em

uma proteína, há a dificuldade de identificar visualmente dentre todos os átomos, aqueles que possuem os parâmetros geométricos para caracterização de uma interação  $\pi$ . Cada um desses vetores constitui uma possível interação  $\pi$  e deve ser analisado individualmente com o objetivo de constatar se o comprimento desse vetor e o seu ângulo com o vetor normal ao plano do anel caracterizam uma interação válida (Figura 2.2). O mapeamento vetorial substitui este trabalho manual com exatidão matemática.



**Figura 2.2** 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 código PDB 1GRA com seus átomos. Estrutura gerada a partir de dados cristalográficos (Karplus e Schulz 1989), visualizada com a utilização do software Discovery Studio 3.1<sup>TM</sup>.

Mesmo com a utilização de softwares para visualização gráfica, face à alta densidade atômica representada pela imagem, a identificação destas interações, além de ser trabalhosa face à necessidade de manualmente serem determinados os centroides e a sua distância ao objeto de destino, não é exata no tocante ao ângulo existente entre o vetor normal ao plano do anel e ao vetor distância da interação. Exemplo desse procedimento é apresentado na Figura 2.3.



**Figura 2.3** Sequência de poses de software de visualização gráfica onde pode ser visto o trabalho de busca de uma possível interação  $\pi$  entre o centroide do resíduo de aminoácido TYR 23 do monômero A de uma glutatona redutase com os átomos adjacentes: **(a)** localização da molécula dotada de anel plano, **(b)** determinação manual do centroide e **(c)** processo intuitivo para a localização de possíveis interações  $\pi$ . Estrutura cristalográfica código PDB 1GRA (Karplus e Schulz, 1989) visualizada utilizando o Discovery Studio 3.1<sup>TM</sup>.



### 2.1.1 O CONCEITO VETORIAL APLICADO ÀS ESTRUTURAS CRISTALOGRÁFICAS

As coordenadas da proteína de código PDB 1GRA foram obtidas do PDB. Uma vez que estas coordenadas fornecem a posição  $x$ ,  $y$  e  $z$  de cada átomo (Tabela 2.1), é possível obter vetorialmente as posições dos centroides dos anéis aromáticos com hibridização  $sp^2$ , os vetores normais aos planos formados por estes anéis, calcular o vetor distância e seu módulo entre o centroide determinado matematicamente e todos os átomos contidos, por fim o ângulo formado pelo vetor normal e todos os vetores distância na estrutura.

**Tabela 2.1** Coordenadas em formato de mol2 do resíduo de aminoácido Tyr 23 da estrutura cristalográfica de uma glutatona redutase código PDB 1GRA utilizada na Figura 2.2 (Karplus e Schulz, 1989)

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

Uma vez efetuados os cálculos, é possível selecionar dentre os resultados obtidos aqueles que venham a atender aos parâmetros de distância e ângulos para as interações  $\pi$  convencionados nas seções 1.2.2, 1.2.3, 1.2.4 e 1.2.5, os quais são agrupados na Tabela 2.2.

**Tabela 2.2** Parâmetros geométricos para as interações  $\pi$ , por tipo

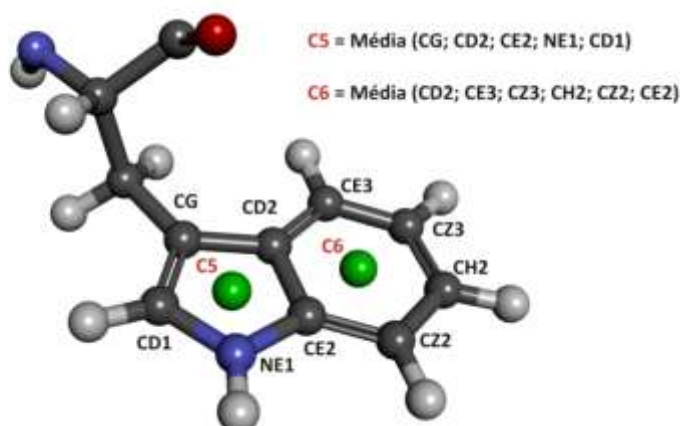
Tipo de Interação	Módulo do vetor distância (Å)	Ângulo entre o vetor normal e o vetor distância (°)
$\pi \cdots \pi$	$3,0 \leq d \leq 4,5$	$\alpha$ e $\beta \leq 30$
C-H $\cdots\pi$	$2,5 \leq d \leq 3,8$	$\alpha \leq 30$
N-H $\cdots\pi$	$2,5 \leq d \leq 3,8$	$\alpha \leq 30$
O-H $\cdots\pi$	$2,5 \leq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

### 2.1.2 CÁLCULO DA COORDENADA DO CENTROIDE

O cálculo da coordenada do centroide é fundamental para que o mapeamento vetorial das interações  $\pi$  seja efetuado, pois é a partir dessa coordenada que todos os outros cálculos são efetuados. Para que seja possível obter a coordenada de um centroide, é necessária a identificação dos átomos que compõem os anéis aromáticos da estrutura cristalográfica. Após isso, calcula-se a média para os valores de  $x$ ,  $y$  e  $z$  das coordenadas destes átomos. Na Figura 2.4 é apresentada a localização destes átomos e suas identificações em um resíduo de aminoácido, já com os centroides calculados conforme a Equação 2.1, onde  $C$  é centroide cuja coordenada é dada por  $x_c$ ,  $y_c$  e  $z_c$ , resultante da média das coordenadas  $x_n$ ,  $y_n$  e  $z_n$ , podendo os valores de  $n$  variarem de 1 a 5 ou de 1 a 6, dependendo da quantidade de átomos que compõem o anel.

$$C(x_c; y_c; z_c) = [P_1(x_1; y_1; z_1) + P_2(x_2; y_2; z_2) + P_3(x_3; y_3; z_3) + \dots + P_n(x_n; y_n; z_n)] / n \quad (2.1)$$

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



**Figura 2.4** Localização dos átomos envolvidos no cálculo dos centroides de um resíduo de aminoácido triptofano. Imagem obtida a partir da estrutura cristalográfica código PDB 1GRA (Karplus e Schulz, 1989) visualizada utilizando o Discovery Studio 3.1<sup>TM</sup>.

Uma vez calculadas as coordenadas dos centroides de todos os anéis aromáticos, estas são organizadas conforme mostra a Tabela 2.1, sendo que na coluna “Átomo” são identificados como “CT”. Após essa operação são armazenadas no formato padrão de arquivo utilizado pelos visualizadores gráficos e também em planilha eletrônica. A utilidade de exportar em um formato de arquivo compatível com os visualizadores gráficos consiste na possibilidade de sobrepor graficamente os centroides no gráfico da macromolécula de origem, poupando o tempo envolvido na sua determinação manualmente.

### 2.1.3 CÁLCULO DO VETOR DISTÂNCIA E SEU MÓDULO

Um vetor no espaço é um segmento orientado que possui uma origem e uma extremidade definidas<sup>2</sup>. Assim sendo, um vetor  $\vec{v} = \overrightarrow{AB}$ , possui a sua origem no ponto  $A (x_1; y_1; z_1)$  e extremidade no ponto  $B (x_2; y_2; z_2)$  e pode ser expresso pela Equação 2.2.

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

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

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

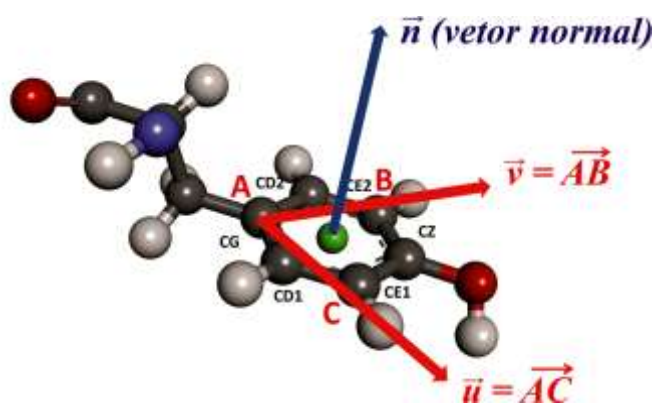
Ao considerar-se que o centroide possui suas coordenadas espaciais e da mesma forma os átomos e outros centroides da estrutura molecular também o possuem, estão então caracterizados os pontos de sua origem  $A (x_1; y_1; z_1)$  e extremidade  $B (x_2; y_2; z_2)$  que darão origem à um vetor. A distância entre estes dois pontos é determinada pelo módulo do vetor  $\vec{v} = \overline{AB}$  e pode ser calculada utilizando-se a Equação 2.3.

$$\|\vec{v}\| = \|\overline{AB}\| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \quad (2.3)$$

Utilizando-se de meios computacionais é possível aplicar a Equação 2.3 para calcular as distâncias entre os centroides e demais átomos e centroides da molécula. Com os parâmetros definidos na Tabela 2.2 é possível a elaboração de um primeiro filtro que permitirá selecionar as distâncias de prováveis interações que possam ocorrer.

#### 2.1.4 CÁLCULO DO VETOR NORMAL AO PLANO DO ANEL AROMÁTICO

Considerando três átomos quaisquer de um anel aromático que formam um plano, é possível determinar os vetores  $\vec{u}$  e  $\vec{v}$ , conforme mostra a Figura 2.5 e o vetor normal  $\vec{n}$  ao plano formado por  $\vec{u}$  e  $\vec{v}$ .



**Figura 2.5** Determinação do vetor normal ao plano do anel aromático de resíduo de aminoácido tirosina. Imagem obtida a partir da estrutura cristalográfica de código PDB 1GRA (Karplus e Schulz, 1989) e visualizada utilizando o Discovery Studio 3.1™.

Uma vez construídos os vetores  $\vec{v} = \overline{AB} = (x_1; y_1; z_1)$  e  $\vec{u} = \overline{AC} = (x_2; y_2; z_2)$  pertencentes ao plano do anel aromático é possível determinar o vetor

normal ao plano ( $\vec{n}$ ) representado pelo determinante de uma matriz quadrada de 3 colunas por 3 linhas. O seu determinante é dado pela Equação 2.4.

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

$$\vec{n} = [(y_1 z_2) - (y_2 z_1)]\vec{i} + [(z_1 x_2) - (z_2 x_1)]\vec{j} + [(x_1 y_2) - (x_2 y_1)]\vec{k} \quad (2.4)$$

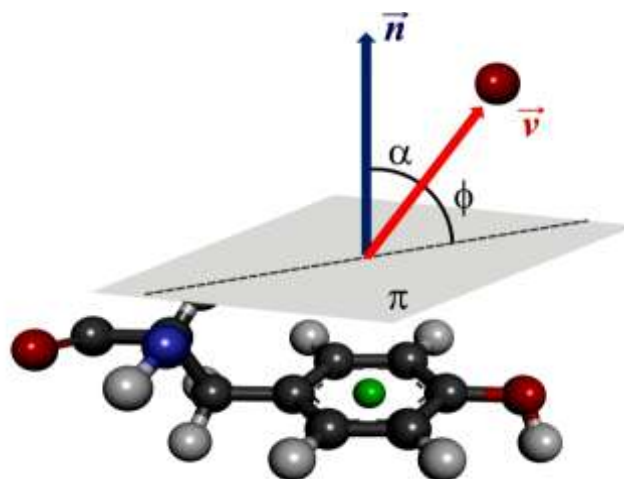
### 2.1.5 CÁLCULO DO ÂNGULO ENTRE O VETOR NORMAL E O VETOR DISTÂNCIA

Havendo um vetor distância  $\vec{v}$  com origem nos centroides determinados pela Equação 2.1 e tendo como destino os átomos e outros centroides da estrutura molecular, existirão tantos vetores distância  $\vec{v}$  quantas combinações centroide-átomo e centroide-centroide existirem, sendo estes vetores  $\vec{v}$  calculados pela Equação 2.2. Tendo como premissa que cada centroide está situado no plano do anel aromático e perpendicular a este plano está o vetor  $\vec{n}$ , conseqüentemente existirão entre os vetores  $\vec{n}$  e  $\vec{v}$  o ângulo  $\alpha$ . Este ângulo, que é o ângulo entre o vetor normal ao plano e o vetor distância do centroide, segundo a Tabela 2.2, é um dos parâmetros geométricos determinante para que uma interação  $\pi$  ocorra.

Steinbruch utiliza a Equação 2.5 para calcular o ângulo  $\alpha$ , que é o ângulo entre o vetor normal ao plano  $\vec{n}$  e o vetor distância  $\vec{v}$ .

$$\sin \alpha = \frac{|\vec{v} \cdot \vec{n}|}{|\vec{v}| \cdot |\vec{n}|}, \text{ assim sendo } \alpha = \sin^{-1} \frac{|\vec{v} \cdot \vec{n}|}{|\vec{v}| \cdot |\vec{n}|} \quad (2.5)$$

A Figura 2.6 ilustra a aplicação dos conceitos da Equação 2.5 sobre o anel aromático de uma tirosina, utilizando o vetor  $\vec{n}$  calculado e anteriormente apresentado na Figura 2.5.



**Figura 2.6** 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 sobre estrutura cristalográfica de código PDB 1GRA (Karplus e Schulz, 1989) e visualizada utilizando o Discovery Studio 3.1<sup>TM</sup>.

A relação de ângulos  $\alpha$  obtidos conforme relatado acima permite a criação de um segundo filtro, selecionando os ângulos válidos a existência de uma interação  $\pi$ , segundo a Tabela 2.2.

## 2.2 UTILIZAÇÃO DE PLANILHA ELETRÔNICA PARA CÁLCULO VETORIAL

Estabelecida 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 operações matemáticas necessárias, uma vez que a quantidade de combinações centroide-átomo e centroide-centroide afetam geometricamente o volume dos cálculos envolvidos.

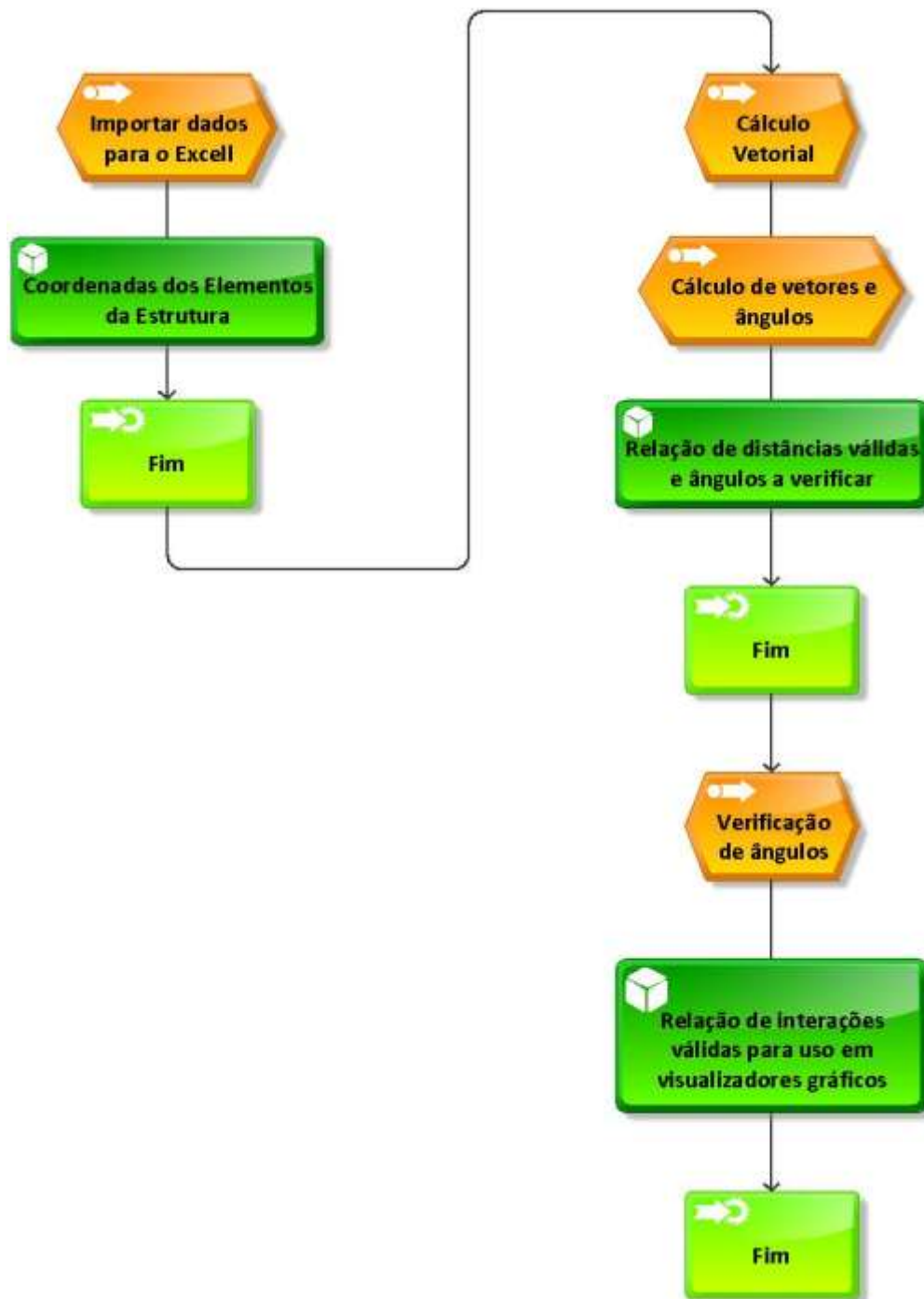
Utilizou-se a planilha eletrônica Excel da suíte Microsoft Office desenvolvida pela Microsoft (Microsoft.Com, 2012), para a realização de todos os cálculos necessários. A adoção deste software baseou-se além da capacidade de cálculos matemáticos,

também na possibilidade de importar os dados contidos nos arquivos do PDB e após o processamento, exportar os resultados em formatos de arquivos compatíveis com os utilizados pelos visualizadores gráficos. Assim sendo, criou-se uma planilha padrão com diversas pastas nas quais cada uma das Equações adotadas possui sua fórmula aplicada nas células pertinentes. Como resultado, obteve-se uma ferramenta capaz de processar o mapeamento de interações  $\pi$  em proteínas e DNA.

Embora esse experimento tenha sido efetuado tanto em proteínas quanto em DNA, optou-se pela utilização do DNA como base para este trabalho por possuir uma estrutura molecular menor. Foram selecionadas três estruturas cristalográficas de DNA, possuindo estas a mesma quantidade de pares de bases e estas apresentando a mesma sequência, de forma que comparações geométricas possam ser efetuadas entre elas. As estruturas escolhidas apresentam como único diferencial a presença ou não de ligantes e o seu posicionamento, caso presentes. Assim há uma estrutura de DNA sem ligante na estrutura obtida do PDB de código 1BNA (Drew et al., 1981) convencionada neste trabalho como DNA – SL, um DNA com ligante no sulco da estrutura obtida do PDB de código do 1VZK (Mallena et al., 2004) e aqui denominado DNA – LS e a terceira com ligante intercalado à estrutura obtida do PDB de código do 1G3X (Malinina et al., 2002) que é o DNA – LI.

Como diversos cálculos são necessários e o relacionamento entre estes ocorre de forma sistêmica, onde segundo Turban (Turban, Mclean, Wetherbe, 2004) um sistema coleta, processa, armazena analisa e dissemina informações com um determinado objetivo dentro de um contexto e que conforme Batista (2004) para cada processo de um sistema, há a entrada de dados, seu processamento e como output informações que alimentarão o processo seguinte. De acordo com Sacco e colaboradores (Sacco et al., 2012), a modelagem de processos constitui importante ferramenta para a análise do fluxo de informação pelos processos que formarão um sistema e para obter-se esse fluxo no mapeamento vetorial para as interações  $\pi$ , efetuou-se a modelagem do fluxo das operações matemáticas utilizando-se o software ARIS (ARIS Express, 2009), o qual permite identificar todas as fases do processamento e também a forma de

relacionamento e interdependência entre estas, incluindo testes lógicos e condicionais, obtendo-se uma visão do processamento dos cálculos envolvidos (Figura 2.7).



**Figura 2.7** Fluxograma das operações envolvidas no mapeamento vetorial para as interações  $\pi$ , processo modelado com a utilização do software ARIS (ARIS Express, 2009).



### 2.3 CÁLCULOS UTILIZANDO PLANILHA ELETRÔNICA

Uma vez que as estruturas cristalográficas depositadas no PDB possuem a padronização da sequência dos átomos disposta em função da sua posição na cadeia polipeptídica, conforme mencionado na Seção 2.1.1, é possível a inserção de Equações em uma planilha eletrônica onde os parâmetros dos cálculos são obtidos em posições específicas e utilizados pelas células onde as Equações são transcritas. Assim, para cada aminoácido, base de DNA ou ligante que possui um anel aromático, é possível a construção de uma planilha na qual uma vez inserida a sequência de átomos da estrutura com anel aromático, automaticamente pode-se obter as coordenadas  $x$ ,  $y$  e  $z$  dos centroides e o vetor normal ao plano do anel aromático. Na Figura 2.8 pode ser vista a planilha para o cálculo e a localização dos campos relacionados à coordenada do centroide e o vetor normal ao plano do anel.

Posição	ÁTOMO	X	Y	Z	Base	Bases	Vetor normal
1	CT	17,8785	28,8882	24,0888	C1	0,1722	0,0247 -2,0636 -4,4565
2	OS*	20,568	34,843	23,895	C3	-0,1834	
3	C5*	19,804	33,450	23,850	C3	-0,1903	
4	C4*	20,742	32,288	23,784	C3	0,1732	
5	O4*	20,104	31,123	24,385	C3	-0,3485	
6	C3*	21,109	31,859	22,389	C3	0,1249	
7	O3*	22,485	31,507	22,287	C3	-0,4776	
8	C2*	20,203	30,890	22,078	C3	0,087	
9	C1*	20,065	30,951	23,448	C3	0,2381	
10	N1	18,807	29,371	23,754	Nam	-0,2325	
11	C2	18,897	28,103	24,354	C2	-0,3511	
12	O2	20,033	27,854	24,571	C2	-0,2455	
13	N3	17,758	27,445	24,653	N2	-0,1761	
14	C4	16,568	28,002	24,383	C2	0,1302	
15	N4	15,473	27,302	24,702	N3	-0,088	
16	C5	16,469	29,290	23,779	C2	0,0846	
17	O5	17,592	29,924	23,490	C2	0,0972	
18	H5*	19,900	33,366	23,760	H	0	
19	H5*1	18,938	33,375	23,194	H	0	
20	H5*2	19,403	33,472	24,930	H	0	
21	H4*	21,577	32,495	24,347	H	0	
22	H3*	20,892	32,694	21,880	H	0	
23	H2*1	20,557	30,006	21,272	H	0	
24	H2*2	19,208	31,080	21,780	H	0	
25	H1*	20,921	29,360	23,830	H	0	
26	H41	14,588	27,726	24,487	H	0	
27	H42	15,541	28,380	25,134	H	0	
28	H5	15,466	29,740	23,562	H	0	
29	H6	17,548	30,923	23,025	H	0	

Ponto	X	Y	Z
Ponto 1	18,8076	29,3710	23,7540
Ponto 2	16,4496	29,2900	23,7790
Ponto 3	17,7580	27,4450	24,6530
Vetor normal Centróide 1			
i	0,0247	-2,0636	-4,4565
j			
k			

**Figura 2.8** Imagem da planilha eletrônica construída para o cálculo da coordenada do centroide da base Citosina, 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 2.1 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 2.4 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.

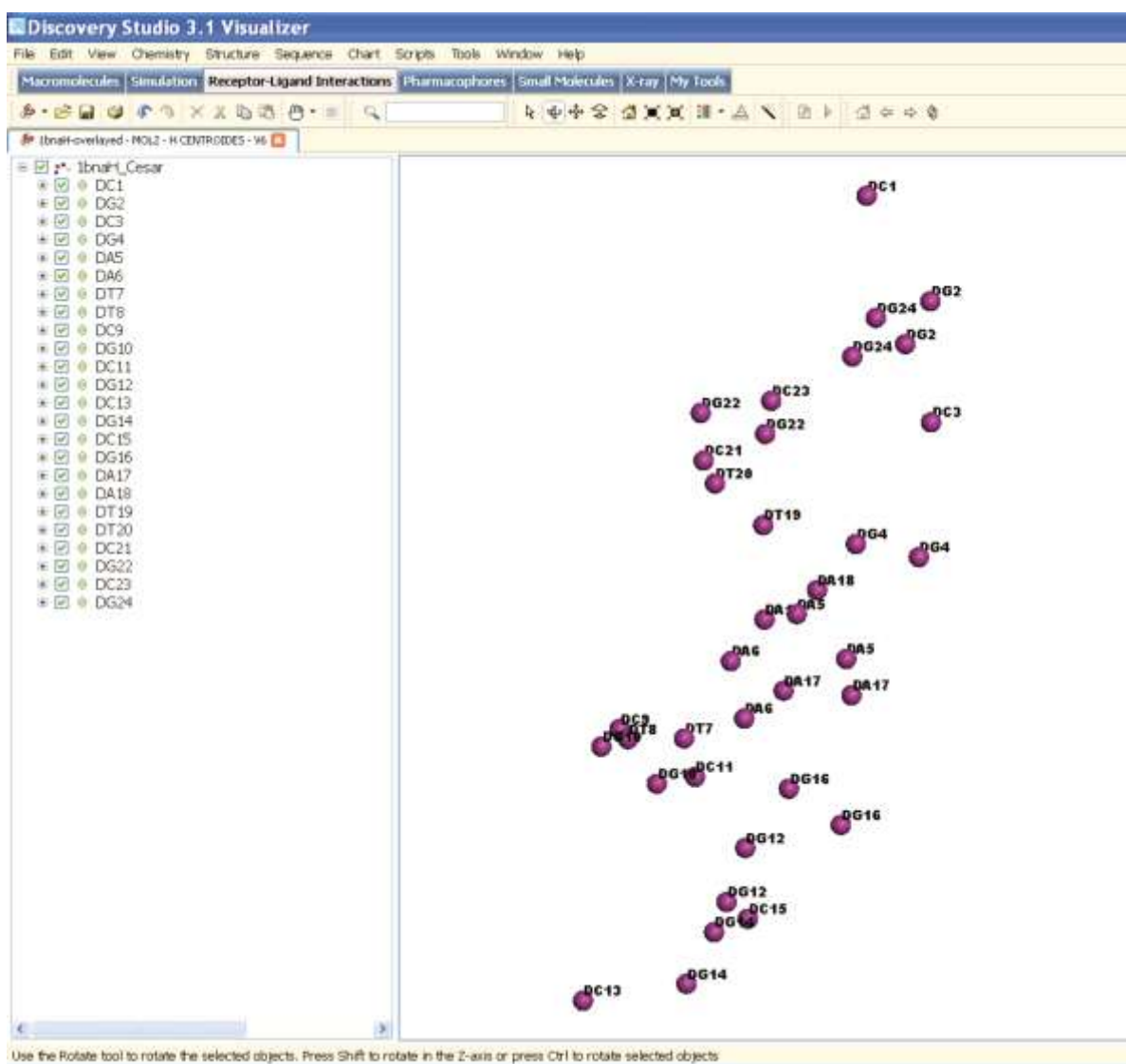
A planilha da Figura 2.8 foi utilizada para o cálculo das coordenadas de todos os centroides das bases citosina e outras semelhantes foram elaboradas para as demais bases (adenina, timina e guanina), bem como para os ligantes existentes nas estruturas de DNA selecionadas.

Os cálculos são efetuados individualmente, base por base e os resultados obtidos das coordenadas dos centroides e vetores normais são relacionados em uma planilha específica (Figura 2.9), que agrupa as informações relativas a todos os centroides da estrutura estudada e utilizados para o cálculo do módulo dos vetores distância e ângulos entre os vetores normais e vetores distância envolvendo todos os centroides determinados e demais átomos da molécula. Esta planilha é um banco de dados unidimensional, sendo que cada linha possui as informações referentes a um único centroide e a coluna “A” representa a identificação do *array*.

	A	B	C	D	E	F	G	H	I	J	K	L	M
		Posição	ÁTOMO	X	Y	Z		Base		Vetor normal			
1	1	1	CT	16.9860	28.9525	24.7440	N.am	1 DC1	0.0000	0.0454	-2.2596	-4.2639	
2	1	2	CT	19.0686	26.6630	21.8522	N.am	2 DG2	0.0000	0.3194	1.0450	2.6454	
3	2	2	CT	18.2408	24.9478	22.6538	N.am	2 DG2	0.0000	0.4479	1.8701	4.4195	
4	3	3	CT	19.2200	23.7527	19.2010	N.am	3 DC3	0.0000	-0.4910	-2.1352	-4.2748	
5	4	4	CT	18.9680	20.3726	16.5868	N.am	4 DG4	0.0000	0.9462	0.9396	2.5163	
6	5	4	CT	17.0217	20.5253	17.2317	N.am	4 DG4	0.0000	1.4840	1.5724	4.2720	
7	6	5	CT	16.8670	17.8856	14.3842	N.am	5 DA5	0.0000	-0.8445	-0.9147	-2.6742	
8	7	5	CT	15.3767	19.3273	14.3915	N.am	5 DA5	0.0000	-1.4510	-1.4682	-4.5489	
9	8	6	CT	13.9228	16.9726	11.6410	N.am	6 DA6	0.0000	-0.8573	-0.7542	-2.7267	
10	9	6	CT	13.5612	18.9813	11.2667	N.am	8 DA6	0.0000	-1.3987	-1.0576	-4.6668	
11	10	7	CT	12.3068	17.6300	8.2173	N.am	7 DT7	0.0000	-1.3687	-0.5625	-4.8095	
12	11	8	CT	10.8322	18.9690	4.6475	N.am	8 DT8	0.0000	-0.7888	-0.2726	-4.7910	
13	12	9	CT	10.8145	20.7042	1.0197	N.am	9 DC9	0.0000	-0.5985	-0.1539	-4.7985	
14	13	10	CT	10.5042	21.6452	-2.9004	N.am	10 DG10	0.0000	0.4260	0.6080	2.7439	
15	14	10	CT	12.1765	20.4530	-2.8932	N.am	10 DG10	0.0000	0.7283	1.0072	4.5993	
16	15	11	CT	13.5883	22.1867	-6.7460	N.am	11 DC11	0.0000	0.7297	-0.3458	-4.7379	
17	16	12	CT	15.3144	21.2526	-10.2524	N.am	12 DG12	0.0000	0.0478	0.6095	2.7860	
18	17	12	CT	14.7003	19.3220	-9.8345	N.am	12 DG12	0.0000	0.1866	0.8898	4.7287	
19	18	13	CT	10.2607	15.7842	-9.1385	N.am	13 DC13	0.0000	-0.3178	0.6288	4.7673	
20	19	14	CT	13.1756	14.9480	-5.4686	N.am	14 DG14	0.0000	-0.5288	-0.3379	-2.7817	
21	20	14	CT	14.0548	16.7860	-5.8425	N.am	14 DG14	0.0000	-0.7753	-0.5364	-4.7525	
22	21	15	CT	14.8450	15.7928	-2.1177	N.am	15 DC15	0.0000	1.1639	1.1775	4.4991	
23	22	16	CT	17.5184	17.7462	0.8080	N.am	16 DG16	0.0000	-0.2423	-1.1447	-2.6165	
24	23	16	CT	15.9770	19.0583	0.4088	N.am	16 DG16	0.0000	-0.4447	-1.7324	-4.4493	
25	24	17	CT	17.7346	21.0458	3.1930	N.am	17 DA17	0.0000	0.1938	1.0724	2.7326	
26	25	17	CT	15.6648	21.1295	3.2950	N.am	17 DA17	0.0000	0.3088	1.7837	4.6814	
27	26	18	CT	16.5374	23.3292	6.1030	N.am	18 DA18	0.0000	0.1348	0.9852	2.7812	
28	27	18	CT	14.8987	22.1570	6.5993	N.am	18 DA18	0.0000	0.2316	1.6529	4.6910	
29	28	19	CT	14.7055	24.1038	9.4935	N.am	19 DT19	0.0000	-0.1742	1.5251	4.5938	
30	29	20	CT	13.0163	24.0010	13.2120	N.am	20 DT20	0.0000	0.0292	1.1504	4.6932	
31	30	21	CT	12.4320	23.2598	17.0342	N.am	21 DC21	0.0000	0.5583	1.0761	4.6807	
32	31	22	CT	12.1062	23.2692	20.9998	N.am	22 DG22	0.0000	-0.6646	-1.1290	-2.5361	
33	32	22	CT	14.0750	22.7160	20.7303	N.am	22 DG22	0.0000	-1.0942	-1.8432	-4.2971	
34	33	23	CT	14.0063	22.2200	24.7978	N.am	23 DC23	0.0000	1.3709	1.4190	4.4089	
35	34	24	CT	16.2776	22.4276	28.0540	N.am	24 DG24	0.0000	-0.7453	-1.1467	-2.4950	
36	35	24	CT	17.0698	24.0698	27.0993	N.am	24 DG24	0.0000	-1.0163	-1.8723	-4.2994	

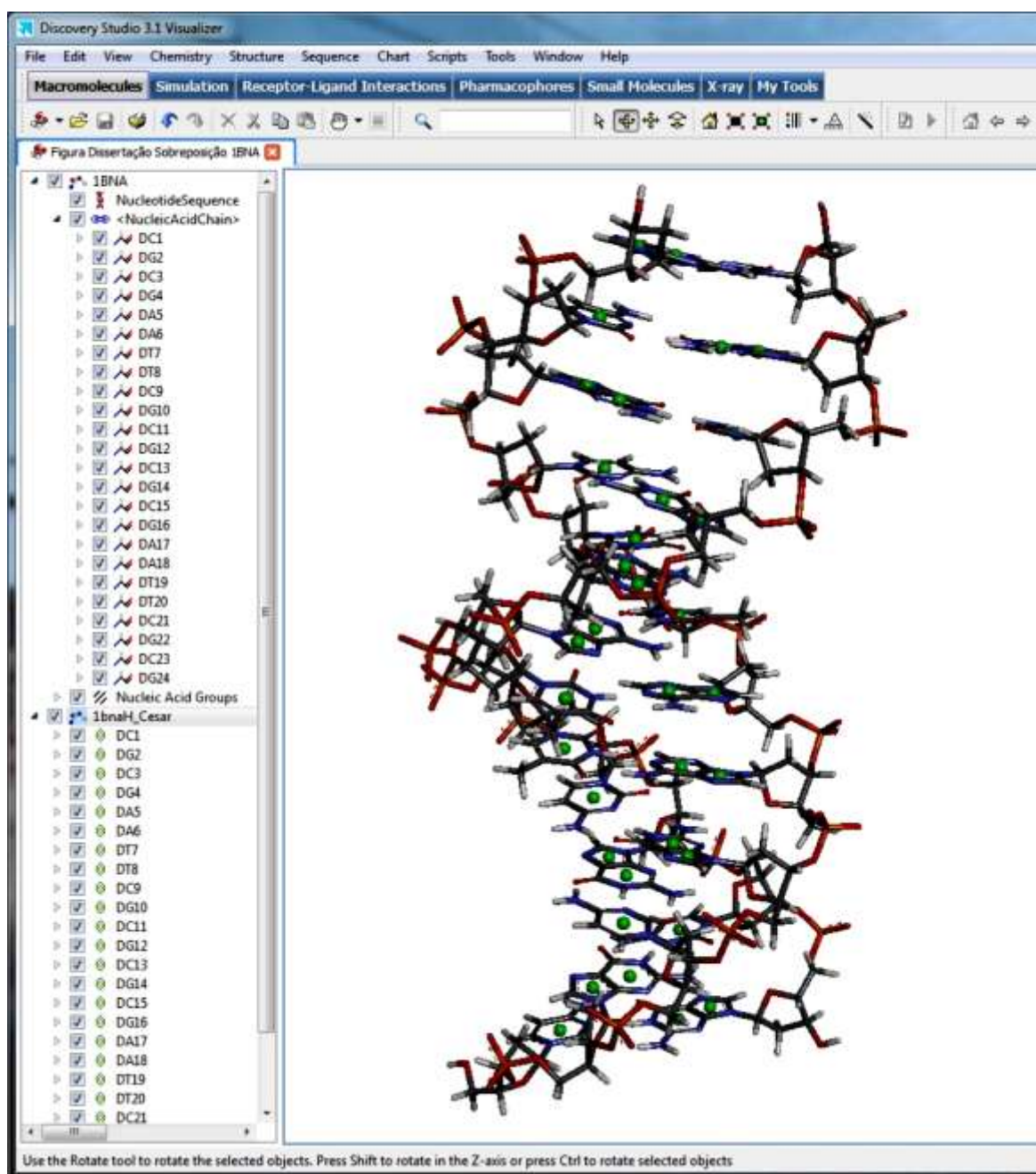
**Figura 2.9** Planilha com a relação das coordenadas dos centroides e vetores normais ao plano de cada anel aromático calculados para o DNA – SL. Utilizada planilha eletrônica do Microsoft Excel.

A relação de centroides obtida e suas respectivas coordenadas é exportada em formato de arquivo compatível aos visualizadores gráficos. O arquivo pode ser utilizado por um visualizador gráfico (Figura 2.10) onde pode ser observada a conformação geométrica do conjunto de centroides na estrutura do DNA.



**Figura 2.10** Visualização dos centroides da estrutura DNA – SL obtidos por cálculo vetorial utilizando planilha eletrônica do Microsoft Excel. Visualização efetuada com a utilização do Discovery Studio 3.1<sup>TM</sup>.

A visualização dos resultados por um software gráfico possibilita a sua inserção de forma sobreposta na visualização da molécula em análise, mostrando de forma precisa e sem intervenção manual, a localização de todos os centroides da estrutura (Figura 2.11).



**Figura 2.11** Visualização dos centroides calculados sobrepostos à estrutura DNA – SL obtidos por cálculo vetorial utilizando planilha eletrônica do Microsoft Excel. Visualização efetuada com a utilização do Discovery Studio 3.1<sup>TM</sup>.

Após o cálculo das coordenadas dos centroides e vetores normais aos planos de cada anel aromático, os dados obtidos são inseridos na planilha 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. A identificação do *array*, fornecida pela coluna “A” do banco de dados dos centroides (Figura 2.8) é o

principal input nessa fase do processo, pois ela retorna aos campos pertinentes os parâmetros a serem utilizados para os cálculos que envolvem as Equações 2.3 e 2.5. O raciocínio adotado foi o de que com os recursos disponíveis na planilha eletrônica torna-se mais prático efetuar todos os cálculos simultaneamente e depois ordenar os resultados em ordem crescente dos valores do módulo do vetor distância, constituindo este o primeiro filtro utilizado para a seleção e avaliação das interações  $\pi$ . A Figura 2.12 reproduz parte desta planilha, identificando suas principais células e funções.

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
1 CT	16,9860	28,9525	24,7440	N.am	1	DC1	0,0000	#DIV/0!	#DIV/0!	#DIV/0!
9 N1	18,0700	29,6610	24,3800	N.am	1	DC1	0,0000	90,0	90,0	90,0
13 C4	15,9170	28,2260	25,1200	C.2	1	DC1	0,0000	-89,9	89,9	89,9
12 N3	17,1430	27,7610	25,3770	N.2	1	DC1	0,0000	90,0	90,0	90,0
10 C2	18,2240	28,4540	25,0150	C.2	1	DC1	0,0000	89,8	89,8	89,8
16 C6	16,8430	30,1710	24,1010	C.2	1	DC1	0,0000	-89,8	89,8	89,8
15 C5	15,7190	29,4420	24,4710	C.2	1	DC1	0,0000	90,0	90,0	90,0
28 H6	16,7403	31,1437	23,5912	H	1	DC1	0,0000	2,49	-89,8	89,8
27 H5	14,7030	29,8146	24,2575	H	1	DC1	0,0000	2,49	89,9	89,9
11 O2	19,3600	28,0140	25,2140	O.2	1	DC1	0,0000	2,60	89,0	89,0
14 N4	14,8280	27,4770	25,4440	N.pl3	1	DC1	0,0000	2,71	88,9	88,9
8 C1'	19,2310	30,4820	23,9440	C.3	1	DC1	0,0000	2,83	89,8	89,8
756 W1	16,8611	26,1006	26,1084	W	24	DC24	0,0000	2,42	80,0	80,0

**Figura 2.12** Imagem da planilha eletrônica construída 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: 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.

O processo é repetido sucessivamente para cada centroide do banco de dados, tendo como resultado um segundo banco de dados no qual constam os valores dos módulos dos vetores distância e o ângulo formado com o vetor normal. Excluem-se dessa relação os valores referentes aos átomos de carbono e demais átomos ou centroides pertencentes à mesma base na qual o centroide está localizado (Figura 2.13).

Como resultado, há uma relação de valores aceitos para o módulo do vetor distância compreendidos entre 2,5 e 4,5 Å, valores estes que são os extremos inferior e superior para as interações  $\pi$  (Tabela 2.2).

O passo seguinte é a seleção dos resultados válidos do módulo do vetor distância em função dos ângulos relativos ao vetor normal e dessa forma são obtidas as possíveis interações  $\pi$  que atendam aos parâmetros de distâncias entre 2,5 e 4,5 Å e ângulos  $\leq 30^\circ$ .

Pos.	X	Y	Z	Aminoácido	Distância	$\alpha$	$ \alpha $
1 CT	16,988	28,9526	24,744	N.am	1 DC1	0	
11 O2	19,36	26,914	25,214	O.2	1 DG1	0	2,59508
14 N4	14,820	27,477	25,444	N.p3	1 DG1	0	2,7003
8 C1'	19,231	30,482	23,944	C.3	1 DG1	0	2,83185
25 H41	14,973	26,5663	25,9194	H	1 DG1	0	3,32154
26 H42	13,8674	27,7992	25,2165	H	1 DG1	0	3,33980
24 H1'	20,1676	29,88	23,8838	H	1 DG1	0	3,42385
4 O4'	19,36	31,583	24,852	O.3	1 DG1	0	3,54561
23 H22	18,842	31,8558	22,795	H	1 DG1	0	3,65279
7 C2'	18,948	31,223	22,647	C.3	1 DG1	0	3,66088
22 H21	18,8003	30,6081	21,7287	H	1 DG1	0	3,68995
42 N7	18,207	27,455	21,836	N.ar	2 DG2	0	3,65965
43 O6	18,083	26,212	22,142	O.ar	2 DG2	0	3,93499
45 O6	15,739	25,916	22,518	O.2	2 DG2	0	3,96616
44 O6	16,904	25,525	22,545	O.2	2 DG2	0	4,07309
2 CT	19,0686	26,683	21,8522	N.am	2 DG2	0	4,23575
44 C8	19,526	27,684	21,429	C.ar	2 DG2	0	4,39473
758 H1	16,8611	26,1906	26,1984	H	24 DG24	0	3,12394
758 H22	19,0789	26,3411	25,7619	H	24 DG24	0	3,49797
742 N1	16,926	25,257	26,804	N.am	24 DG24	0	4,13762
744 N2	19,208	25,386	26,096	N.p3	24 DG24	0	4,41419
795 O	14,095	28,151	21,614	O.3	37 HOH62	0	4,33557
819 H	14,8226	29,3241	21,1139	H	45 HOH77	0	4,24217

**Figura 2.13** Imagem da planilha eletrônica com os resultados válidos do módulo do vetor distância, onde são excluídos os resultados envolvendo átomos e centroides da mesma base, bem como átomos de carbono. Planilha eletrônica do Microsoft Excel.

Ao serem analisados os resultados para ângulos  $\leq 30^\circ$ , pode-se concluir que dentre todos os cálculos efetuados para o centroide da Citosina1, apenas o N7 que ocupa a posição 42 na estrutura apresenta resultados que atendam aos requisitos de uma interação  $\pi$ . Neste caso, há uma interação  $N\cdots\pi$  com um comprimento de 3,65 Å e um ângulo de  $19,07^\circ$ , atendendo ao especificado na Tabela 2.2 (Figura 2.14).

	A	B	C	D	E	F	G	H	I	J	K	L
1	Pos.	CT	X	Y	Z		Base	DC1		Distância	α	[α]
2	1	CT	16,9860	28,9525	24,7440	CT	1	DC1	0,0000			
3	42	N7	18,2070	27,4550	21,6360	N.ar	2	DG2	0,0000	3,6596	19,07	19,07
4	45	O6	15,7390	25,9160	22,5180	O.2	2	DG2	0,0000	3,9662	31,63	
5	2	CT	19,0686	26,6630	21,8522	N.am	2	DG2	0,0000	4,2357	30,58	
6	756	H1	16,8611	26,1906	26,1984	H	24	DG24	0,0000	3,1239	89,87	
7	758	H22	19,0789	26,3411	25,7619	H	24	DG24	0,0000	3,4980	84,37	
8	742	N1	16,9260	25,2570	26,6040	N.am	24	DG24	0,0000	4,1376	88,80	
9	744	N2	19,2080	25,3860	26,0960	N.pl3	24	DG24	0,0000	4,4142	83,54	
10	795	O	14,0950	28,1510	21,6140	O.3	37	HOH62	0,0000	4,3356	44,10	
11	819	H	14,8226	29,3241	21,1139	H	45	HOH77	0,0000	4,2422	44,74	
12												
13												
14												
15												
16												
17												

**Figura 2.14** Imagem da planilha eletrônica com os resultados válidos para o ângulo relativo ao vetor normal. Planilha eletrônica do Microsoft Excel.

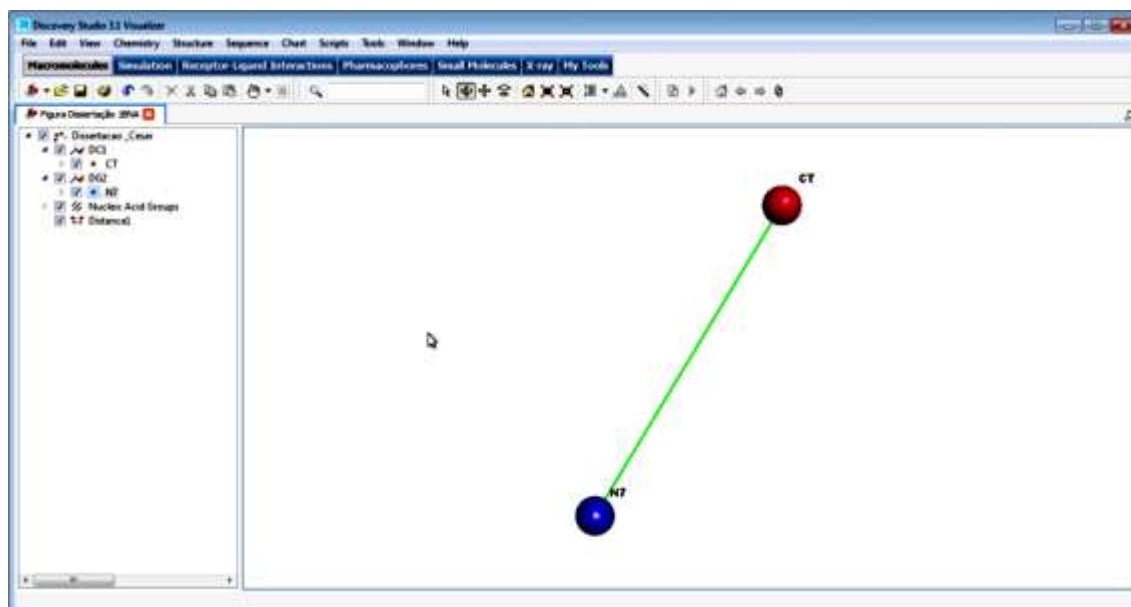
A análise dos ângulos válidos é efetuada sucessivamente para todas as relações de comprimentos de interações centroide-átomo e centroide-centroide válidas. O resultado desse último filtro é uma planilha estruturada na qual figuram as informações sobre o centroide e os elementos com os quais ele possui uma interação  $\pi$ . Para que possa ser exportado para um formato de arquivo reconhecível por um visualizador gráfico, a planilha deve conter um cabeçalho padrão, que pode ser editado com informações que atendam à necessidade do usuário e salva em formato de texto separado por tabulações. Esse procedimento pode ser observado na Figura 2.15, onde estão identificados:

- 1- Campo com os dados do centroide;
- 2- Átomo envolvido na interação  $\pi$  com o centroide;
- 3- Cabeçalho do texto utilizado pelo visualizador gráfico como identificação de arquivo de formato compatível;
- 4- Dados de origem do arquivo. No visualizador gráfico é inserido automaticamente.

#	Created by:	Discovery Studio	Studio	F	G	H	I	J	K	L	M
#	Creation time:	Wed	October	15	16:38:20	E	South	America	Standard	Time	2012
@<TRIPOS>MOLECULE											
Dissertacao_Cesar											
	802	863	25	0	0						
NUCLEIC_ACID											
USER_CHARGES											
@<TRIPOS>ATOM											
1	CT	16.9860	28.9525	24.7440	CT	1	DC1	0.0000			
42	N7	18.2070	27.4550	21.6360	N.ar	2	DG2	0.0000			

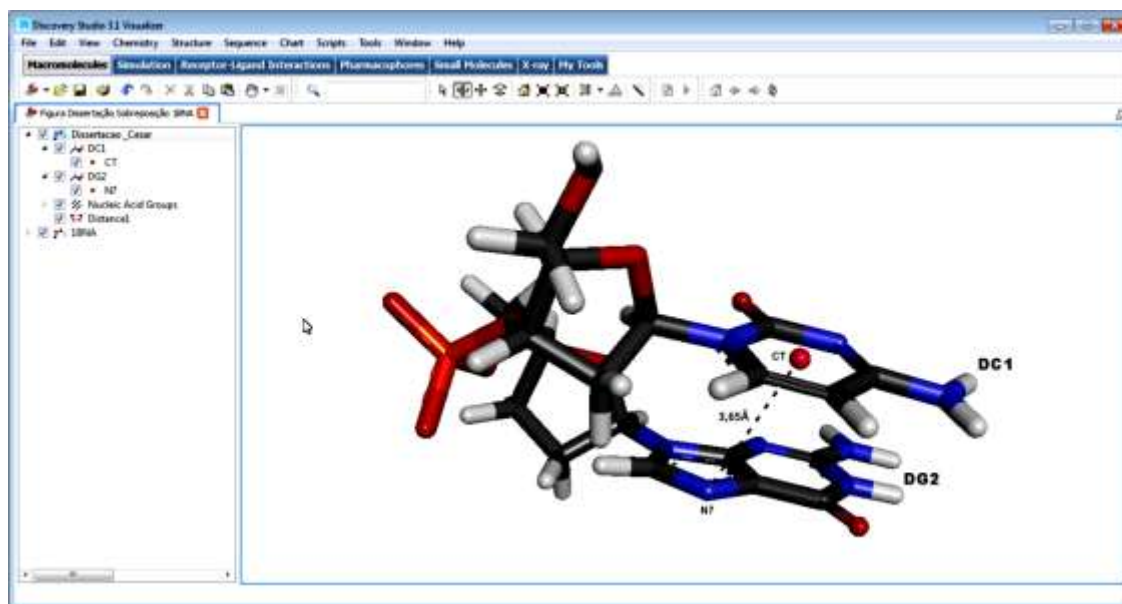
**Figura 2.15** Imagem da planilha eletrônica utilizada para exportar os resultados válidos de uma interação  $\pi$ , em arquivo compatível para visualização em software gráfico. Planilha eletrônica do Microsoft Excel.

A visualização em software gráfico da interação  $N^{\pi}$  obtida por meio de cálculo vetorial pode ser observada na Figura 2.16, onde foram adicionadas as legendas e linha da interação com o auxílio do Discovery Studio 3.1. A sobreposição dessa interação ao arquivo do DNA sem ligante pode ser visualizada na Figura 2.17.



**Figura 2.16** Visualização da imagem da interação  $N^{\pi}$  obtida por meio de cálculo vetorial na estrutura DNA – SL DNA. Utilizado o software Discovery Studio 3.1™.





**Figura 2.17** Visualização da interação N... $\pi$  obtida por meio de cálculo vetorial na estrutura DNA – SL DNA. Utilizado o software Discovery Studio 3.1<sup>TM</sup>.

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 a todas as interações ficarem sobrepostas a macromolécula, substituindo o trabalho manual por uma rotina computacional caracterizada por sua exatidão.

## 2.4 CONCLUSÕES DO CAPÍTULO 2

Utilizando a álgebra linear, foram construídas fórmulas no Microsoft Excel para o processamento das coordenadas de biomoléculas disponíveis de forma padronizada no PDB. Os exemplos citados acima mostram a exatidão dos cálculos obtidos pela ferramenta computacional desenvolvida para o mapeamento vetorial, atingindo o objetivo proposto por este trabalho. No capítulo seguinte, a escolha de três estruturas de DNA com a mesma sequência de nucleotídeos permitiu identificar outras aplicações imediatas e inovadoras do mapeamento vetorial, aplicáveis também em proteínas.

# Capítulo 3

---

RESULTADOS E DISCUSSÕES

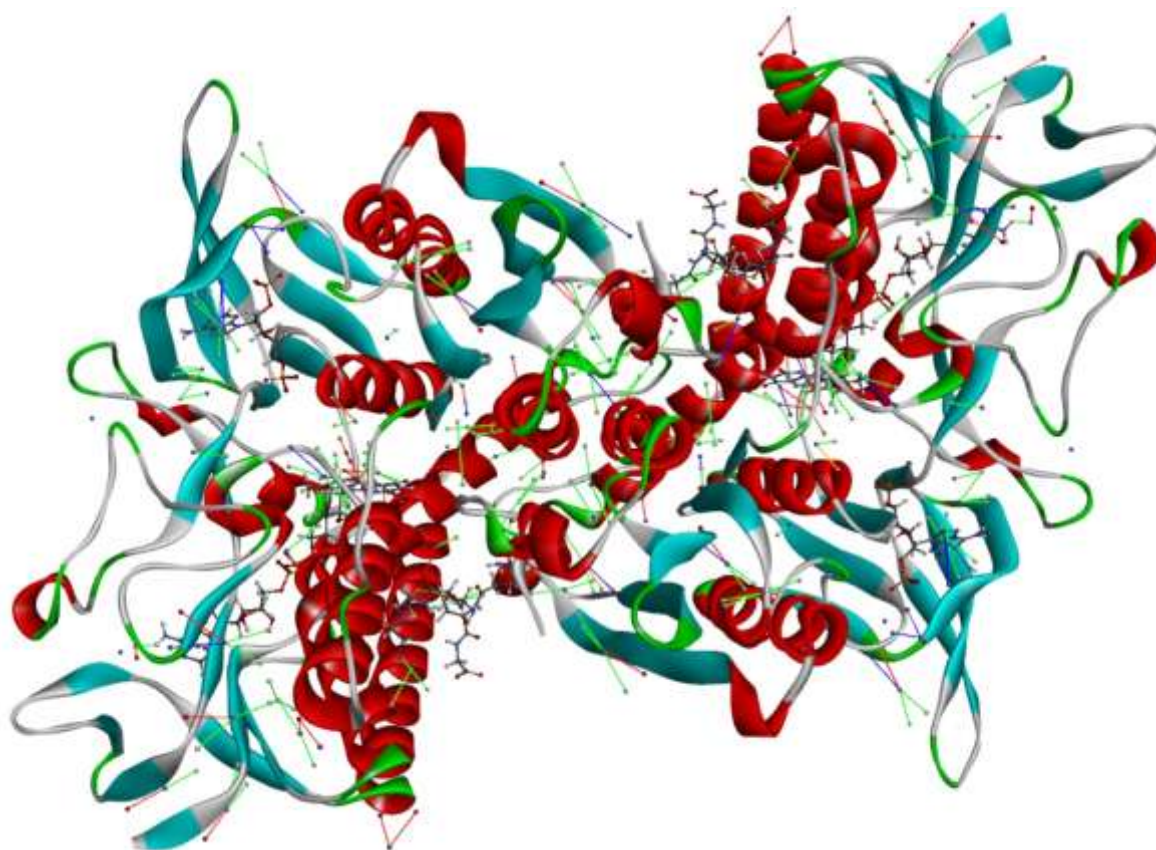
Inicialmente os conceitos descritos no Capítulo 2 (Materiais e Métodos) foram aplicados no mapeamento de uma proteína código PDB 1GRA (Anexo 3.1) e em uma estrutura de DNA com 12 pares de base código PDB 1BNA. Ao comparar os resultados, observou-se que embora a quantidade de interações para a proteína seja em maior que as identificadas no DNA, esta apresenta uma menor relação de interações por átomos (268 interações em 14500 átomos) do que a estrutura de DNA (164 interações em 996 átomos). Outra premissa para a utilização das estruturas de DNA é que na proteína (PDB ID 1GRA) a distribuição das interações  $\pi$  ocorre de forma dispersa em sua estrutura (Figura 3.1).

**Tabela 3.1** Comparativo de interações  $\pi$  por tipo: proteína e DNA

Tipo de Interação	Proteína PDB 1GRA		DNA PDB 1BNA	
	$N \cdots \pi$	22	8,2%	88
$O \cdots \pi$	38	14,2%	25	15,2%
$S \cdots \pi$	4	1,5%	-	-
$\pi \cdots \pi$	-	-	32	19,2%
$C - H \cdots \pi$	192	71,6%	5	3,0%
$N - H \cdots \pi$	12	4,5%	14	8,5%
Totais de interações por estrutura	268	100%	164	100%

Em função disso, foram selecionadas três estruturas de DNA para que as aplicações do mapeamento vetorial fossem conhecidas com uma maior abrangência. As estruturas tridimensionais de DNA utilizadas neste trabalho apresentam mesma sequência de nucleotídeos e foram denominadas da seguinte maneira:

- DNA – SL → DNA sem ligante na estrutura obtida do PDB de código 1BNA (Drew et al., 1981).
- DNA – LS → DNA com ligante no sulco da estrutura obtida do PDB de código do 1VZK (Mallena et al., 2004).
- DNA – LI → DNA com ligante intercalado à estrutura obtida do PDB de código do 1G3X (Malinina et al., 2002).



**Figura 3.1** Representação gráfica das interações  $\pi$  mapeadas vetorialmente para a proteína código PDB 1GRA sobrepostas à sua estrutura. Estrutura gerada a partir de dados cristalográficos (Karplus e Schulz 1989), visualizada com a utilização do software Discovery Studio 3.1<sup>TM</sup>.

São apresentadas neste capítulo aplicações imediatas e inovadoras do mapeamento vetorial para as análises qualitativas e quantitativas das possíveis interações  $\pi$  existentes em três estruturas cristalográficas de DNA com mesma sequência de nucleotídeos e que apresentam como único diferencial a presença ou não de ligantes e o seu posicionamento, estando estes presentes um de forma intercalada ortogonalmente ao eixo da hélice e o outro alinhado ao sulco menor. Inicialmente, após o mapeamento das interações  $\pi$ , são obtidos para cada estrutura mapeada os tipos de interações presentes e sua quantidade, possibilitando uma avaliação se a presença de ligantes interfere qualitativamente ou quantitativamente nos tipos de interações existentes. Possuindo estas informações e com os resultados do mapeamento vetorial das três estruturas de DNA exportados em arquivo digital compatível com os visualizadores

gráficos, será mostrado em imagens como o caminho eletrostático das interações  $\pi$  é afetado pela presença dos ligantes. Outro ponto em destaque está nas imagens da distorção causada pelos ligantes sobrepondo as estruturas com e sem ligantes, e pela primeira vez mensuradas as distorções existentes base por base e analisando estatisticamente o grau de deformação resultante em função da presença de ligantes. Por fim, é efetuada uma análise por tipo de interação  $\pi$ , considerando as variáveis módulo do vetor distância das interações e vetor normal ao plano do anel aromático e ângulos entre o vetor distância e o vetor normal ao plano do anel aromático. Todas as análises comparativas efetuadas possuem como padrão de referência o DNA que não possui ligante em sua estrutura cristalográfica.

### 3.1 ANÁLISES QUALITATIVAS E QUANTITATIVAS DE INTERAÇÕES $\pi$

Segundo Janiak (2000) e Wheeler e colaboradores (2010), as interações  $\pi$  são responsáveis pela conformação do DNA, e de acordo com Nechipurenko e colaboradores (2008), a presença de ligantes altera as propriedades físico-químicas dos ácidos desoxirribonucléicos, o que vem a alterar as interações  $\pi$  qualitativa e quantitativamente.

#### 3.1.1 VARIAÇÃO DAS INTERAÇÕES $\pi$ EM FUNÇÃO DA PRESENÇA DE LIGANTES

Com base na metodologia desenvolvida para o mapeamento vetorial de macromoléculas, por meio de cálculo vetorial foram identificadas as interações  $\pi$  presentes nas estruturas de DNA sem ligante na estrutura obtida do PDB de código 1BNA (Drew et al., 1981), na estrutura com ligante no sulco da estrutura obtida do PDB de código do 1VZK (Mallena et al., 2004) e por fim na estrutura com ligante intercalado à estrutura obtida do PDB de código do 1G3X (Malinina et al., 2002).

Os cálculos efetuados para o DNA sem ligante (DNA – SL), nos quais foram determinadas as possíveis interações  $\pi$  encontram-se no Anexo 3.2 e a análise dos ângulos no Anexo 3.3. Para o DNA com o ligante no sulco (DNA – LS), as possíveis

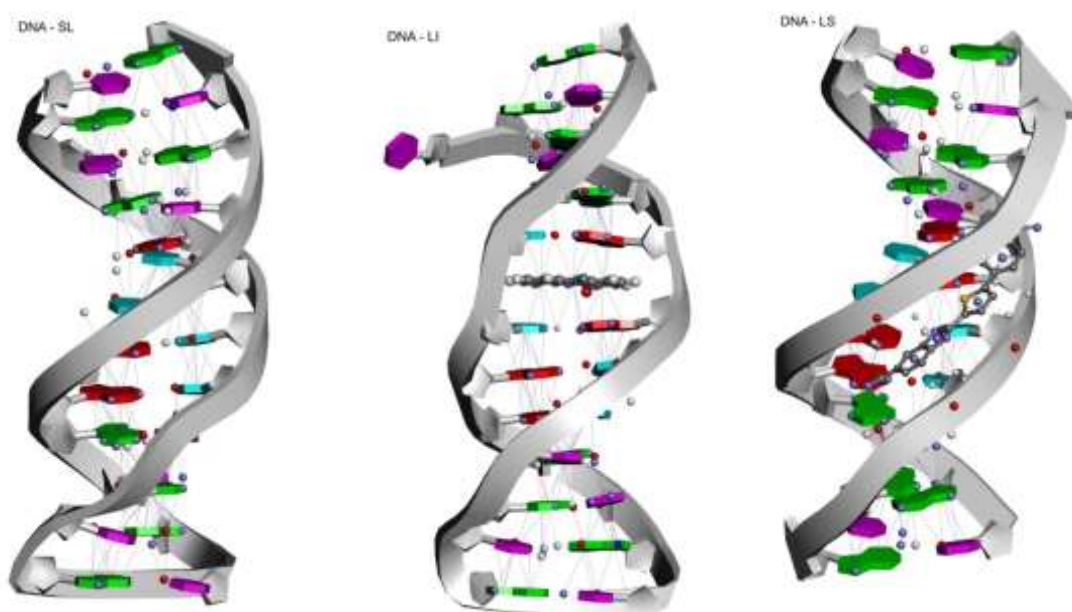
interações  $\pi$  encontram-se no Anexo 3.4 e a análise dos ângulos no Anexo 3.5. Finalizando, os cálculos para o DNA com o ligante intercalado (DNA – LI) encontram-se no Anexo 3.6 e a respectiva análise pelos ângulos válidos no Anexo 3.7. Sendo conhecidas as interações válidas para as três estruturas de DNA obtidas pelo filtro final dos ângulos efetuado respectivamente pelas análises compreendidas nos Anexos 3.3, 3.5 e 3.7, foram comparadas as interações  $\pi$  identificadas por meio de análises qualitativas e quantitativas do Anexo 3.8.

O resultado dessas análises foi efetuado considerando-se os tipos de interações existentes em valores absolutos, incluindo as existentes entre as bases e os ligantes (Tabela 3.2).

**Tabela 3.2** Interações  $\pi$  identificadas por estrutura de DNA, incluídas as interações presentes nos ligantes. Valores absolutos das quantidades por tipo de interação

Tipo de interação $\pi$	1BNA DNA – SL	1G3X DNA – LI	1VZK 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

As interações quantificadas na Tabela 3.2 podem ser visualizadas na Figura 3.2 e os arquivos para uso em visualizadores gráficos encontram-se no Anexo 3.9.



**Figura 3.2** Imagens das estruturas de DNA com as interações  $\pi$  identificadas por mapeamento vetorial.

Os resultados expressos na Tabela 3.2 podem ser expressos em valores relativos por tipo de interação, permitindo observar a participação proporcional de cada tipo de interação presente nas estruturas.

**Tabela 3.3** Interações  $\pi$  identificadas por estrutura de DNA, incluídas as interações presentes nos ligantes. Valores relativos das quantidades por tipo de interação

Tipo de interação $\pi$	1BNA DNA – SL	1G3X DNA – LI	1VZK 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%

Em função dos valores relativos da Tabela 3.3, pode-se observar que tomando como padrão os valores relativos ao DNA sem ligante e comparando-os com o que possui ligante intercalado, não existe uma variação expressiva nos diversos tipos de interações

identificadas. Por outro lado, uma comparação do DNA que possui o ligante no sulco com o DNA sem ligante, há uma variação nas proporções das interações  $N^{\cdots}\pi$ ,  $O^{\cdots}\pi$  e  $C-H^{\cdots}\pi$ .

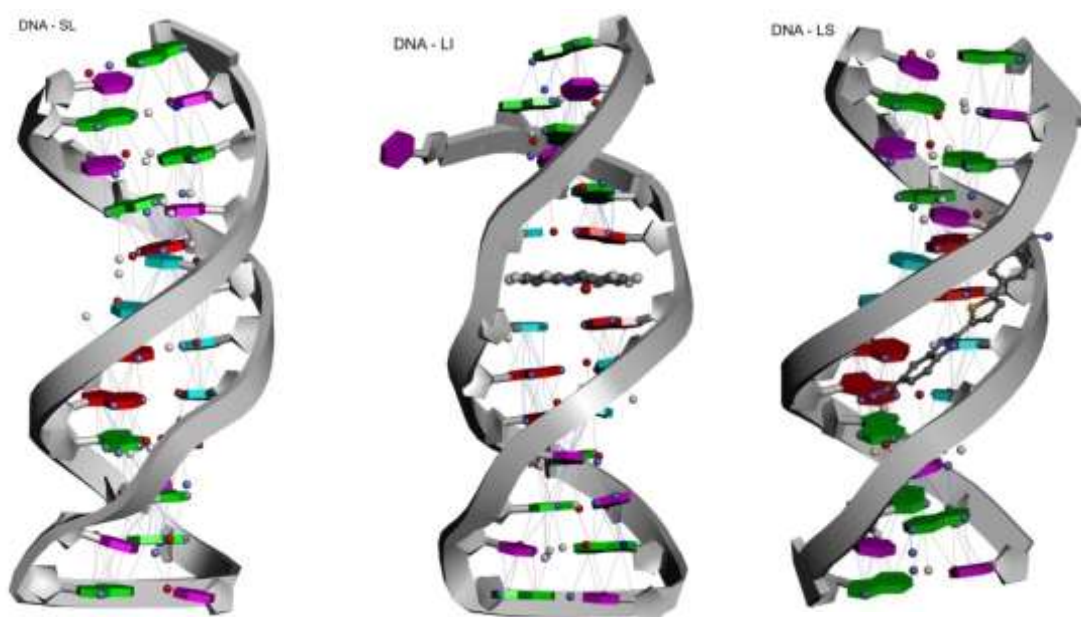
Repetindo o procedimento anterior, porém excluindo da análise as interações  $\pi$  entre as bases e os ligantes e mantendo as interações  $\pi$  existentes somente entre as bases, foi construída a Tabela 3.4. Nesta tabela é notada a redução das interações totais remanescentes entre as bases para as estruturas que possuem ligante.

**Tabela 3.4** Interações  $\pi$  identificadas por estrutura de DNA, não computadas as interações presentes nos ligantes. Valores absolutos das quantidades por tipo de interação

Tipo de interação $\pi$	1BNA DNA – SL	1G3X DNA – LI	1VZK DNA – LS
$N^{\cdots}\pi$	88	75	81
$O^{\cdots}\pi$	25	23	24
$\pi^{\cdots}\pi$	32	25	31
$C - H^{\cdots}\pi$	5	5	5
$N - H^{\cdots}\pi$	14	9	13
Totais de interações por estrutura	164	137	154

As interações da Tabela 3.4 identificadas nas estruturas estudadas podem ser visualizadas com software gráfico e os arquivos digitais referentes a estas estão contidos no Anexo 3.10. Com base nestes arquivos, foram criadas poses referentes às estruturas de DNA estudados com as interações  $\pi$  identificadas, porém, sem a presença dos ligantes. Pode-se observar que para o DNA com o ligante intercalado é criada uma lacuna na região do ligante interrompendo a cadeia de interações  $\pi$ , o que não ocorre com o DNA com o ligante no sulco (Figura 3.3).





**Figura 3.3** Imagens das estruturas de DNA com as interações  $\pi$  identificadas por mapeamento vetorial, excluindo as interações entre o ligante e a base.

Embora existam variações em termos absolutos nas quantidades de interações ao excluirmos as existentes entre as bases e os ligantes conforme demonstrado na Tabela 3.4 e visualizado na Figura 3.3, em valores relativos por tipo de interação, é observada que a proporcionalidade entre os tipos de interações  $\pi$  nas bases é mantida (Tabela 3.5).

**Tabela 3.5** Interações  $\pi$  identificadas por estrutura de DNA, não computadas as interações presentes nos ligantes. Valores relativos das quantidades por tipo de interação

Tipo de interação $\pi$	1BNA DNA – SL	1G3X DNA – LI	1VZK DNA – LS
N $\cdots\pi$	54%	55%	53%
O $\cdots\pi$	15%	17%	16%
$\pi\cdots\pi$	20%	18%	20%
C - H $\cdots\pi$	3%	4%	3%
N - H $\cdots\pi$	9%	7%	8%
Totais de interações por estrutura	100%	100%	100%

Comparando os valores absolutos da Tabela 3.2 com os da Tabela 3.4 é possível quantificar as interações  $\pi$  exclusivas entre bases e ligante (Tabela 3.6).

**Tabela 3.6** Interações  $\pi$  identificadas por estrutura de DNA existentes entre as bases e ligantes. Valores absolutos das quantidades por tipo de interação, tomando como referência a estrutura de DNA sem ligante

Tipo de interação $\pi$	1BNA DNA – SL	1G3X DNA – LI	1VZK DNA – LS
N $\cdots\pi$	0	8	1
O $\cdots\pi$	0	0	8
$\pi\cdots\pi$	0	8	0
C - H $\cdots\pi$	0	4	13
N - H $\cdots\pi$	0	4	0
Totais de interações por estrutura	0	24	22

### 3.1.2 MAPEAMENTO DAS VARIAÇÕES DAS INTERAÇÕES $\pi$

Embora a quantidade total de interações não possua uma variação expressiva, o tipo de posicionamento do ligante apresenta uma significativa alteração no tipo de interação existente entre base e ligante.

Utilizando os dados obtidos e tabulados no Anexo 3.8 e com os recursos da tabela dinâmica existentes na planilha eletrônica Excel, é 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 por base e também a sua posição nesta, o que pode ser observado nas Tabelas de 3.7 a 3.42.

**Tabela 3.7** Interações  $\pi$  com origem na base C1 com os átomos e centroides relacionados por bases com o respectivo comprimento  $d$  da interação e ângulo  $\alpha$

Centroide de origem da interação: C1								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA DNA – SL	
			$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)
G2	CT6	$\pi\cdots\pi$			4,15	28,46		
	N7	N $\cdots\pi$			3,66	14,99	3,66	19,07

Analisando as informações da Tabela 3.7, é possível verificar que o centroide da citosina C1 possui uma interação  $\pi \cdots \pi$  com o centroide do anel aromático com seis átomos da guanina G2 no comprimento de 4,15 Å para o DNA com ligante no sulco. Para as outras duas estruturas de DNA, essa interação não existe. Para as estruturas com ligante no sulco e sem ligante, o mesmo centroide da base C1 possui uma interação  $N \cdots \pi$  com o nitrogênio N7 da mesma guanina G2 e em ambos os casos, embora coincidentemente o comprimento das interações seja de 3,66 Å e o ângulo  $\alpha$  apresenta uma variação de 4,08°. O mesmo raciocínio aplica-se às tabelas subsequentes até a Tabela 3.42, onde pode ser constatado por meio dos valores do comprimento da interação e ângulo  $\alpha$  que a presença do ligante e a forma como este se posiciona na estrutura, afeta as propriedades físico-químicas destas (Nechipurenko et al., 2008).

**Tabela 3.8** 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  $\alpha$

Centroide de origem da interação: G2-A								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA 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 3.8 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 altera o comprimento desta interação e o valor do ângulo existente entre o vetor normal ao plano do centroide e o vetor distância centroide-átomo.

**Tabela 3.9** Interações  $\pi$  com origem na base G2, anel aromático com 5 átomos, com os átomos e centroides relacionados por bases com o respectivo comprimento  $d$  da interação e ângulo  $\alpha$ 

Centroide de origem da interação: G2-B								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA DNA – SL	
			$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)
C1	O2	O $\cdots$ $\pi$			4,03	24,00	4,15	28,62
C3	N3	N $\cdots$ $\pi$	3,51	19,73	3,52	3,86	3,50	2,54
	N4	N $\cdots$ $\pi$	4,01	14,10				
	CT6	$\pi$ $\cdots$ $\pi$	4,12	25,08	3,81	25,02	3,78	20,69
G24	N1	N $\cdots$ $\pi$					4,17	29,88
	N3	N $\cdots$ $\pi$	4,03	28,51				
	N2	N $\cdots$ $\pi$	3,14	8,10	3,65	26,11	3,60	18,49
	H21	N - H $\cdots$ $\pi$			3,52	14,27		
G22	H1	N - H $\cdots$ $\pi$			3,71	27,12	3,80	28,72

**Tabela 3.10** Interações  $\pi$  com origem na base C3, com os átomos e centroides relacionados por bases com o respectivo comprimento  $d$  da interação e ângulo  $\alpha$ 

Centroide de origem da interação: C3								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA DNA – SL	
			$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)
G2	CT6	$\pi$ $\cdots$ $\pi$			3,86	18,79	3,94	22,59
	CT5	$\pi$ $\cdots$ $\pi$			3,81	23,63	3,78	22,14
	N3	N $\cdots$ $\pi$	3,75	24,70	3,59	21,96	3,63	16,83
	N9	N $\cdots$ $\pi$			3,75	18,45	3,88	21,70
	H2'2	C - H $\cdots$ $\pi$	2,88	26,29				
G4	N7	N $\cdots$ $\pi$	3,95	25,77	3,86	12,35	3,92	12,74
	CT6	$\pi$ $\cdots$ $\pi$	4,48	27,35	4,32	24,53	4,28	25,77
	O6	O $\cdots$ $\pi$	3,39	23,07				

**Tabela 3.11** Interações  $\pi$  com origem na base G4, anel aromático com 6 átomos, com os átomos e centroides relacionados por bases com o respectivo comprimento  $d$  da interação e ângulo  $\alpha$ 

Centroide de origem da interação: G4-A								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA DNA – SL	
			$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)
C3	O2	O $\cdots$ $\pi$	4,32	18,45	4,05	21,88	3,97	23,78
	N1	N $\cdots$ $\pi$			4,15	25,44	4,06	19,17
A5	CT6	$\pi$ $\cdots$ $\pi$	3,56	19,53	3,92	29,57	3,93	28,18
	N7	N $\cdots$ $\pi$	3,23	10,13	3,61	14,56	3,53	15,83
	H62	N - H $\cdots$ $\pi$					3,72	29,01

**Tabela 3.12** Interações  $\pi$  com origem na base G4, anel aromático com 5 átomos, com os átomos e centroides relacionados por bases com o respectivo comprimento  $d$  da interação e ângulo  $\alpha$ 

Centroide de origem da interação: G4-B								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA DNA – SL	
			$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)
C3	O2	O $\cdots$ $\pi$	4,28	17,21	3,80	8,84	3,70	9,07
A5	N1	N $\cdots$ $\pi$	3,55	24,57	3,65	26,81	3,73	24,41
	CT6	$\pi$ $\cdots$ $\pi$	3,82	28,64	3,83	27,00	3,89	27,22
	CT5	$\pi$ $\cdots$ $\pi$	3,34	5,75	3,39	14,04	3,49	10,33
	N3	N $\cdots$ $\pi$	3,87	25,14			4,02	29,46
G22	H22	N - H $\cdots$ $\pi$	3,11	14,22			3,55	20,75
	H21	N - H $\cdots$ $\pi$			3,52	19,43		

**Tabela 3.13** Interações  $\pi$  com origem na base A5, anel aromático com 6 átomos, com os átomos e centroides relacionados por bases com o respectivo comprimento  $d$  da interação e ângulo  $\alpha$ 

Centroide de origem da interação: A5-A								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA DNA – SL	
			$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)
9AC700	H6	C - H $\cdots$ $\pi$	3,50	24,24				
	H5	C - H $\cdots$ $\pi$	3,52	19,90				
	CT	$\pi$ $\cdots$ $\pi$	3,78	26,73				
A6	CT6	$\pi$ $\cdots$ $\pi$			3,93	23,19	4,13	28,89
	N7	N $\cdots$ $\pi$			3,58	5,68	3,72	12,78
G4	CT5	$\pi$ $\cdots$ $\pi$			3,83	29,28	3,89	27,58
	CT6	$\pi$ $\cdots$ $\pi$	3,56	27,98	3,92	26,87		
	N3	N $\cdots$ $\pi$	3,68	21,06	3,52	18,88	3,58	13,61
	N9	N $\cdots$ $\pi$	3,46	19,90	3,76	20,75	3,75	25,56

**Tabela 3.14** Interações  $\pi$  com origem na base A5, anel aromático com 5 átomos, com os átomos e centroides relacionados por bases com o respectivo comprimento  $d$  da interação e ângulo  $\alpha$ 

Centroide de origem da interação: A5-B								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA DNA – SL	
			$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)
9AC700	CT5	$\pi \cdots \pi$	3,82	21,17				
	CT	$\pi \cdots \pi$	3,54	18,61				
	N10	$N \cdots \pi$	3,88	24,42				
A6	CT5	$\pi \cdots \pi$			3,56	16,61	3,64	16,80
	CT6	$\pi \cdots \pi$			3,77	16,75	3,90	21,94
	N3	$N \cdots \pi$			4,02	28,66		
	N7	$N \cdots \pi$			3,97	26,29	4,02	25,50
	N9	$N \cdots \pi$			4,20	27,10		
G4	CT5	$\pi \cdots \pi$	3,34	11,31	3,39	9,80	3,49	12,11
	N3	$N \cdots \pi$	3,62	17,77	3,58	21,58	3,73	21,42
	N1	$N \cdots \pi$	3,50	21,41	3,43	18,53	3,55	16,27
	N2	$N \cdots \pi$	4,09	29,47			4,00	29,68

**Tabela 3.15** Interações  $\pi$  com origem na base A6, anel aromático com 6 átomos, com os átomos e centroides relacionados por bases com o respectivo comprimento  $d$  da interação e ângulo  $\alpha$ 

Centroide de origem da interação: A6-A								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA DNA – SL	
			$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)
9AC700	CT	$\pi \cdots \pi$	3,60	22,33				
	H7	$N - H \cdots \pi$	3,39	21,23				
	H6	$C - H \cdots \pi$	3,59	28,27				
A5	CT5	$\pi \cdots \pi$			3,77	23,56	3,90	24,18
	CT6	$\pi \cdots \pi$			3,93	26,34	4,13	28,77
	N3	$N \cdots \pi$			3,75	13,66	3,67	11,16
	N9	$N \cdots \pi$			3,99	22,82	4,06	25,16
T7	CT6	$\pi \cdots \pi$			3,74	22,42	3,84	25,82
	H73	$C - H \cdots \pi$	2,90	28,23				
	O4	$O \cdots \pi$	4,29	27,70				

**Tabela 3.16** Interações  $\pi$  com origem na base A6, anel aromático com 5 átomos, com os átomos e centroides relacionados por bases com o respectivo comprimento  $d$  da interação e ângulo  $\alpha$ 

Centroide de origem da interação: A6-B								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA DNA – SL	
			$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)
9AC700	CT6	$\pi \cdots \pi$	3,98	28,11				
	CT6	$\pi \cdots \pi$	3,48	16,72				
A5	CT6	$\pi \cdots \pi$			3,56	13,35	3,64	14,87
	N3	$N \cdots \pi$			4,00	24,34	3,88	21,21
	N1	$N \cdots \pi$			3,43	12,91	3,57	16,89
	H2	$C - H \cdots \pi$			3,93	25,48	3,79	22,68
T7	CT6	$\pi \cdots \pi$	3,53	17,16	3,60	16,11	3,55	11,48
	N3	$N \cdots \pi$	3,55	14,72	3,37	5,80	3,45	11,29
	O4	$O \cdots \pi$	4,12	23,30				
	N1	$N \cdots \pi$			4,05	29,35	3,85	26,12
	H3	$N - H \cdots \pi$					3,71	26,93

**Tabela 3.17** Interações  $\pi$  com origem na base T7, com os átomos e centroides relacionados por bases com o respectivo comprimento  $d$  da interação e ângulo  $\alpha$ 

Centroide de origem da interação: T7								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA DNA – SL	
			$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)
A6	CT6	$\pi \cdots \pi$	4,09	27,25	3,74	22,41	3,84	18,67
	N3	$N \cdots \pi$	3,40	15,39	3,64	17,86	3,68	21,48
	N1	$N \cdots \pi$	3,97	28,70				
	CT5	$\pi \cdots \pi$	3,53	10,22	3,60	20,68	3,55	16,67
	N9	$N \cdots \pi$			3,85	22,66	4,06	23,65
T8	CT6	$\pi \cdots \pi$			4,00	26,67	4,09	26,44
	H72	$C - H \cdots \pi$	2,94	24,17			3,45	29,29
	O4	$O \cdots \pi$	3,61	29,08	3,78	27,52	3,77	24,92

**Tabela 3.18** Interações  $\pi$  com origem na base T8, com os átomos e centroides relacionados por bases com o respectivo comprimento  $d$  da interação e ângulo  $\alpha$ 

Centroide de origem da interação: T8								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA DNA – SL	
			$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)
C9	CT6	$\pi \cdots \pi$	3,96	15,37				
	N3	$N \cdots \pi$	4,31	24,06				
	N4	$N \cdots \pi$	3,99	23,87	3,69	18,96	3,72	21,91
	H5	$C - H \cdots \pi$			3,64	24,04		
T7	CT6	$\pi \cdots \pi$	4,05	29,58	4,00	27,59	4,09	25,78
	N3	$N \cdots \pi$			3,81	26,43	3,95	27,33
	N1	$N \cdots \pi$	4,07	18,26	4,03	23,00	4,06	19,54
	O2	$O \cdots \pi$	3,51	16,71	3,69	12,56	3,71	14,18

**Tabela 3.19** Interações  $\pi$  com origem na base C9, com os átomos e centroides relacionados por bases com o respectivo comprimento  $d$  da interação e ângulo  $\alpha$ 

Centroide de origem da interação: C9								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA DNA – SL	
			$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)
G10	N7	$N \cdots \pi$	3,82	13,55	3,90	12,93	3,86	13,75
	CT6	$\pi \cdots \pi$	4,06	24,07	4,08	18,97	4,04	15,49
	CT5	$\pi \cdots \pi$					4,15	26,32
T8	CT6	$\pi \cdots \pi$	3,96	23,33			4,02	28,17
	N3	$N \cdots \pi$	3,95	14,66	3,88	25,93	3,85	25,34
	O2	$O \cdots \pi$	4,19	17,73	3,47	9,28	3,50	11,67
	N1	$N \cdots \pi$	4,07	27,45			3,96	25,87

**Tabela 3.20** Interações  $\pi$  com origem na base G10, anel aromático com 6 átomos, com os átomos e centroides relacionados por bases com o respectivo comprimento  $d$  da interação e ângulo  $\alpha$ 

Centroide de origem da interação: G10-A								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA DNA – SL	
			$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)
C9	N1	$N \cdots \pi$	3,85	12,42	3,92	11,30	3,93	10,86
	CT6	$\pi \cdots \pi$			4,08	28,05	4,04	26,16
	O2	$O \cdots \pi$	3,92	23,67	4,19	25,37	4,26	22,24
C11	N4	$N \cdots \pi$	3,96	22,46	3,67	19,32	4,21	21,20
	H41	$N - H \cdots \pi$	3,78	8,37				



**Tabela 3.21** Interações  $\pi$  com origem na base G10, anel aromático com 5 átomos, com os átomos e centroides relacionados por bases com o respectivo comprimento  $d$  da interação e ângulo  $\alpha$ 

Centroide de origem da interação: G10-B								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA DNA – SL	
			$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)
C9	CT6	$\pi \cdots \pi$			4,12	29,18	4,15	29,23
	O2	O $\cdots$ $\pi$	3,74	16,80	3,92	14,81	4,12	16,79
	N3	N $\cdots$ $\pi$	3,58	21,64	3,68	19,87	3,77	17,68
C11	N4	N $\cdots$ $\pi$	3,76	12,97	3,70	20,56	4,05	14,04
	H42	N - H $\cdots$ $\pi$			3,56	7,22		
G14	O6	O $\cdots$ $\pi$			3,49	29,16	3,63	27,97
	H1	N - H $\cdots$ $\pi$	3,58	29,94				
G16	H1	N - H $\cdots$ $\pi$					3,74	29,76
	H22	N - H $\cdots$ $\pi$					3,48	28,53
	H21	N - H $\cdots$ $\pi$	3,35	24,45				

**Tabela 3.22** Interações  $\pi$  com origem na base C11, com os átomos e centroides relacionados por bases com o respectivo comprimento  $d$  da interação e ângulo  $\alpha$ 

Centroide de origem da interação: C11								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA DNA – SL	
			$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)
G10	N3	N $\cdots$ $\pi$	3,33	9,82	3,28	23,28	3,58	14,04
	CT5	$\pi \cdots \pi$	4,13	28,28			4,45	29,23
	N9	N $\cdots$ $\pi$					4,49	28,50
G12	CT6	$\pi \cdots \pi$	3,96	22,63	4,30	23,58	4,02	19,58
	N7	N $\cdots$ $\pi$	3,72	16,08	3,93	14,29	3,73	3,37

**Tabela 3.23** Interações  $\pi$  com origem na base G12, anel aromático com 6 átomos, com os átomos e centroides relacionados por bases com o respectivo comprimento  $d$  da interação e ângulo  $\alpha$ 

Centroide de origem da interação: G12-A								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA DNA – SL	
			$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)
C11	CT6	$\pi \cdots \pi$					4,02	26,52
	O2	O $\cdots$ $\pi$	3,83	14,72	4,09	14,44	3,99	15,49
	N1	N $\cdots$ $\pi$	3,84	19,93	4,11	21,67	4,06	16,51

**Tabela 3.24** Interações  $\pi$  com origem na base G12, anel aromático com 5 átomos, com os átomos e centroides relacionados por bases com o respectivo comprimento  $d$  da interação e ângulo  $\alpha$ 

Centroide de origem da interação: G12-B								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA DNA – SL	
			$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)
C11	O2	O $\cdots$ $\pi$	4,06	24,18	4,13	16,12	4,16	22,47
	N3	N $\cdots$ $\pi$	3,42	15,58	3,97	25,40	3,68	22,84
G14	H21	N - H $\cdots$ $\pi$	3,67	24,68	3,54	19,87	3,70	23,02

**Tabela 3.25** Interações  $\pi$  com origem na base C13, com os átomos e centroides relacionados por bases com o respectivo comprimento  $d$  da interação e ângulo  $\alpha$ 

Centroide de origem da interação: C13								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA DNA – SL	
			$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)
G14	N7	N $\cdots$ $\pi$			3,48	16,19		
	O6	O $\cdots$ $\pi$	4,15	23,99				

**Tabela 3.26** Interações  $\pi$  com origem na base G14, anel aromático com 6 átomos, com os átomos e centroides relacionados por bases com o respectivo comprimento  $d$  da interação e ângulo  $\alpha$ 

Centroide de origem da interação: G14-A								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA DNA – SL	
			$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)
C13	O2	O $\cdots$ $\pi$	3,87	20,67	3,44	11,02	3,49	9,36
	N1	N $\cdots$ $\pi$			3,81	27,07	4,37	25,89
	H1'	C - H $\cdots$ $\pi$	3,72	14,80				
C15	N3	N $\cdots$ $\pi$	3,65	25,46	3,91	26,02	3,93	25,24
	CT6	$\pi$ $\cdots$ $\pi$	3,59	13,96	3,77	19,79	3,84	16,57
	N4	N $\cdots$ $\pi$	3,95	26,01	4,12	22,03	4,17	25,52

**Tabela 3.27** Interações  $\pi$  com origem na base G14, anel aromático com 5 átomos, com os átomos e centroides relacionados por bases com o respectivo comprimento  $d$  da interação e ângulo  $\alpha$ 

Centroide de origem da interação: G14-B								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA DNA – SL	
			$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)
C13	O2	O $\cdots$ $\pi$	3,74	13,90	3,68	23,33	3,78	24,65
C15	O2	O $\cdots$ $\pi$			3,68	29,10	3,73	28,19
	N3	N $\cdots$ $\pi$	3,33	9,28	3,54	7,68	3,59	7,50
	CT6	$\pi$ $\cdots$ $\pi$	3,86	25,61	3,83	22,15	3,94	21,15
G10	H1	N - H $\cdots$ $\pi$	3,53	28,61	3,62	29,95	3,73	29,89
G12	N2	N $\cdots$ $\pi$	3,98	29,71			3,76	28,37
	H21	N - H $\cdots$ $\pi$			3,39	20,92		

**Tabela 3.28** Interações  $\pi$  com origem na base C15, com os átomos e centroides relacionados por bases com o respectivo comprimento  $d$  da interação e ângulo  $\alpha$ 

Centroide de origem da interação: C15								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA DNA – SL	
			$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)
G14	N7	N $\cdots$ $\pi$	4,03	27,34	4,29	28,02	4,29	27,27
	CT6	$\pi$ $\cdots$ $\pi$	3,59	11,42	3,77	12,88	3,84	11,74
	N9	N $\cdots$ $\pi$	3,63	7,76	3,77	13,99	3,88	11,04
	N3	N $\cdots$ $\pi$			3,78	27,81	3,92	29,07
	CT5	$\pi$ $\cdots$ $\pi$			3,83	26,84	3,94	28,94
G16	N7	N $\cdots$ $\pi$	4,15	24,01	4,20	17,81	4,18	18,32
	O6	O $\cdots$ $\pi$	3,55	23,50			3,83	27,53
	CT6	$\pi$ $\cdots$ $\pi$					4,42	28,73

**Tabela 3.29** Interações  $\pi$  com origem na base G16, anel aromático com 6 átomos, com os átomos e centroides relacionados por bases com o respectivo comprimento  $d$  da interação e ângulo  $\alpha$ 

Centroide de origem da interação: G16-A								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA DNA – SL	
			$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)
C15	O2	O $\cdots$ $\pi$	3,99	18,99	4,05	28,52	4,00	28,68
	N1	N $\cdots$ $\pi$			4,21	21,48	4,07	17,13
A17	N7	N $\cdots$ $\pi$	3,59	24,70	3,41	20,56	3,62	15,81
	H62	N - H $\cdots$ $\pi$	3,56	23,23	3,69	26,51		

**Tabela 3.30** Interações  $\pi$  com origem na base G16, anel aromático com 5 átomos, com os átomos e centroides relacionados por bases com o respectivo comprimento  $d$  da interação e ângulo  $\alpha$ 

Centroide de origem da interação: G16-B								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA DNA – SL	
			$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)
C15	O2	O $\cdots$ $\pi$	3,86	12,05	3,56	2,17	3,58	8,41
	N1	N $\cdots$ $\pi$					4,40	29,51
	N3	N $\cdots$ $\pi$					3,82	27,27
A17	N7	N $\cdots$ $\pi$					3,94	27,94
	N1	N $\cdots$ $\pi$	3,68	26,33	3,75	26,53	3,82	28,60
	CT5	$\pi$ $\cdots$ $\pi$	3,55	22,11	3,42	15,08	3,57	17,71
	CT6	$\pi$ $\cdots$ $\pi$			3,65	28,42	3,85	25,74
	N6	N $\cdots$ $\pi$	3,62	24,91				
G10	H22	N - H $\cdots$ $\pi$	3,54	25,11			3,58	24,96
	H21	N - H $\cdots$ $\pi$			3,56	23,73		

**Tabela 3.31** Interações  $\pi$  com origem na base A17, anel aromático com 6 átomos, com os átomos e centroides relacionados por bases com o respectivo comprimento  $d$  da interação e ângulo  $\alpha$ 

Centroide de origem da interação: A17-A								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA DNA – SL	
			$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)
A18	CT6	$\pi$ $\cdots$ $\pi$	3,73	25,12	3,92	29,33	3,89	27,25
	N7	N $\cdots$ $\pi$	3,41	8,36	3,59	12,08	3,51	9,79
G16	CT5	$\pi$ $\cdots$ $\pi$			3,65	26,04	3,85	27,01
	N3	N $\cdots$ $\pi$	3,38	13,62	3,27	9,52	3,52	10,48
	N9	N $\cdots$ $\pi$					3,95	28,11

**Tabela 3.32** Interações  $\pi$  com origem na base A17, anel aromático com 5 átomos, com os átomos e centroides relacionados por bases com o respectivo comprimento  $d$  da interação e ângulo  $\alpha$ 

Centroide de origem da interação: A17-B								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA DNA – SL	
			$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)
A18	CT5	$\pi \cdots \pi$	3,47	12,06	3,43	18,53	3,54	16,40
	CT6	$\pi \cdots \pi$	3,68	23,07	3,67	21,15	3,67	19,83
	N3	$N \cdots \pi$	3,88	28,50				
	N7	$N \cdots \pi$					3,86	26,39
	N1	$N \cdots \pi$	3,87	28,73				
G16	CT5	$\pi \cdots \pi$	3,55	21,97	3,42	16,28	3,57	17,08
	N3	$N \cdots \pi$	3,53	21,33	3,64	27,74	3,76	23,04
	N1	$N \cdots \pi$	3,58	23,02	3,41	12,33	3,51	16,24
	N2	$N \cdots \pi$	3,56	21,87	3,78	29,53	3,73	26,76
	H1	$N - H \cdots \pi$			3,73	25,38	3,79	27,91

**Tabela 3.33** Interações  $\pi$  com origem na base A18, anel aromático com 6 átomos, com os átomos e centroides relacionados por bases com o respectivo comprimento  $d$  da interação e ângulo  $\alpha$ 

Centroide de origem da interação: A18-A								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA DNA – SL	
			$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)
A17	CT5	$\pi \cdots \pi$	3,68	22,47	3,67	24,44	3,67	21,46
	CT6	$\pi \cdots \pi$	3,73	24,73	3,92	25,69	3,89	27,54
	N3	$N \cdots \pi$	3,53	16,49	3,48	15,42	3,53	12,17
	N9	$N \cdots \pi$	3,66	23,35	3,88	22,86	3,88	26,01
T19	CT6	$\pi \cdots \pi$	3,94	26,87	3,80	27,56		
	H72	$C - H \cdots \pi$	3,10	29,13			3,10	25,51

**Tabela 3.34** Interações  $\pi$  com origem na base A18, anel aromático com 5 átomos, com os átomos e centroides relacionados por bases com o respectivo comprimento  $d$  da interação e ângulo  $\alpha$ 

Centroide de origem da interação: A18-B								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA DNA – SL	
			$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)
A17	CT5	$\pi \cdots \pi$	3,47	12,56	3,43	13,05	3,54	15,30
	N3	$N \cdots \pi$	3,82	27,26	3,74	26,45	3,91	28,02
	N1	$N \cdots \pi$	3,46	11,17	3,31	12,99	3,43	9,60
	H2	$C - H \cdots \pi$			3,62	29,42		
T19	CT6	$\pi \cdots \pi$	3,61	12,87	3,52	16,71	3,49	15,64
	N3	$N \cdots \pi$	3,39	9,69	3,36	7,38	3,40	11,70
	N1	$N \cdots \pi$	4,10	28,91				
	H3	$N - H \cdots \pi$			3,58	22,79	3,66	26,06
	O4	$O \cdots \pi$					3,92	28,25

**Tabela 3.35** Interações  $\pi$  com origem na base T19, com os átomos e centroides relacionados por bases com o respectivo comprimento  $d$  da interação e ângulo  $\alpha$ 

Centroide de origem da interação: T19								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA DNA – SL	
			$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)
9AC700	CT6	$\pi \cdots \pi$	3,79	21,68				
A18	CT6	$\pi \cdots \pi$	3,94	24,34	3,80	26,77	3,93	26,25
	CT5	$\pi \cdots \pi$	3,61	19,99	3,52	18,36	3,49	15,59
	N3	$N \cdots \pi$	3,62	14,26	3,46	13,75	3,41	14,05
	N9	$N \cdots \pi$	4,07	23,59	3,87	27,94	4,00	28,63
T20	CT6	$\pi \cdots \pi$			3,90	28,53	4,09	29,52
	H72	$C - H \cdots \pi$					3,08	26,07
	O4	$O \cdots \pi$			3,78	28,75	3,88	28,15
	H73	$C - H \cdots \pi$			3,24	29,48		

**Tabela 3.36** Interações  $\pi$  com origem na base T20, com os átomos e centroides relacionados por bases com o respectivo comprimento  $d$  da interação e ângulo  $\alpha$ 

Centroide de origem da interação: T20								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA 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

**Tabela 3.37** Interações  $\pi$  com origem na base C21, com os átomos e centroides relacionados por bases com o respectivo comprimento  $d$  da interação e ângulo  $\alpha$ 

Centroide de origem da interação: C21								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA DNA – SL	
			$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)
G22	N7	N $\cdots \pi$	4,10	23,32	3,79	13,16	3,73	15,15
	O6	O $\cdots \pi$	3,99	29,95				
	CT6	$\pi \cdots \pi$	4,14	21,39	4,02	14,95	3,98	17,14
	CT5	$\pi \cdots \pi$	3,97	19,37	4,15	27,79	4,08	26,70
	N1	N $\cdots \pi$	4,16	29,39				
T20	CT6	$\pi \cdots \pi$			3,98	29,42	3,94	28,32
	N1	N $\cdots \pi$	4,00	24,33	4,01	26,69	3,99	25,04
	O2	O $\cdots \pi$	3,47	17,50	3,74	10,68	3,72	11,51
	N3	N $\cdots \pi$			3,83	25,30	3,76	25,80

**Tabela 3.38** Interações  $\pi$  com origem na base G22, anel aromático com 6 átomos, com os átomos e centroides relacionados por bases com o respectivo comprimento  $d$  da interação e ângulo  $\alpha$ 

Centroide de origem da interação: G22-A								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA DNA – SL	
			$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)
C21	O2	O $\cdots$ $\pi$	4,43	27,41	4,33	22,24	4,28	23,19
	N1	N $\cdots$ $\pi$	3,86	6,66	3,93	12,77	3,82	12,80
	CT6	$\pi$ $\cdots$ $\pi$	4,14	27,84	4,02	28,75	3,98	29,85
C23	CT6	$\pi$ $\cdots$ $\pi$	3,61	20,89				
	N3	N $\cdots$ $\pi$	3,58	29,28				
	N4	N $\cdots$ $\pi$	3,46	21,44	3,76	11,17	3,78	13,02

**Tabela 3.39** Interações  $\pi$  com origem na base G22, anel aromático com 6 átomos, com os átomos e centroides relacionados por bases com o respectivo comprimento  $d$  da interação e ângulo  $\alpha$ 

Centroide de origem da interação: G22-B								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA DNA – SL	
			$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)	$d$ (Å)	$\alpha$ (°)
C21	O2	O $\cdots$ $\pi$	4,20	20,46	4,11	12,89	4,02	12,39
	N1	N $\cdots$ $\pi$	4,33	28,00				
	CT6	$\pi$ $\cdots$ $\pi$	3,97	22,60				
	N3	N $\cdots$ $\pi$	3,77	11,87	3,74	22,09	3,72	22,61
C23	CT6	$\pi$ $\cdots$ $\pi$	3,72	24,67				
	N3	N $\cdots$ $\pi$	3,13	5,63	3,65	22,73	3,67	20,10
	N4	N $\cdots$ $\pi$			3,92	20,03	3,91	19,91
G2	H1	N - H $\cdots$ $\pi$	3,71	24,30			3,64	26,69
G4	H1	N - H $\cdots$ $\pi$					3,55	29,74
	H22	N - H $\cdots$ $\pi$					3,60	28,72
	H21	N - H $\cdots$ $\pi$			3,53	25,76		



**Tabela 3.40** Interações  $\pi$  com origem na base C23, com os átomos e centroides relacionados por bases com o respectivo comprimento **d** da interação e ângulo  $\alpha$ 

Centroide de origem da interação: C23								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA DNA – SL	
			<b>d</b> (Å)	$\alpha$ (°)	<b>d</b> (Å)	$\alpha$ (°)	<b>d</b> (Å)	$\alpha$ (°)
G22	CT6	$\pi \cdots \pi$	3,61	23,88				
	CT5	$\pi \cdots \pi$			3,95	24,85		
	N3	$N \cdots \pi$	3,62	28,96	3,27	6,52	3,54	14,36
	N9	$N \cdots \pi$	3,56	11,18			4,05	26,85
G24	N7	$N \cdots \pi$	3,49	15,37	3,76	14,67	3,71	10,99
	CT6	$\pi \cdots \pi$	3,75	4,49			3,98	22,32

**Tabela 3.41** Interações  $\pi$  com origem na base G24, anel aromático com 6 átomos, com os átomos e centroides relacionados por bases com o respectivo comprimento **d** da interação e ângulo  $\alpha$ 

Centroide de origem da interação: G24-A								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA DNA – SL	
			<b>d</b> (Å)	$\alpha$ (°)	<b>d</b> (Å)	$\alpha$ (°)	<b>d</b> (Å)	$\alpha$ (°)
C23	O2	$O \cdots \pi$	4,50	13,92	3,64	9,32	3,83	15,09
	N1	$N \cdots \pi$	3,98	19,55	4,14	23,64	3,89	18,28
	CT6	$\pi \cdots \pi$	3,75	24,19	4,17	29,91	3,98	27,26
	N3	$N \cdots \pi$	3,94	22,31				

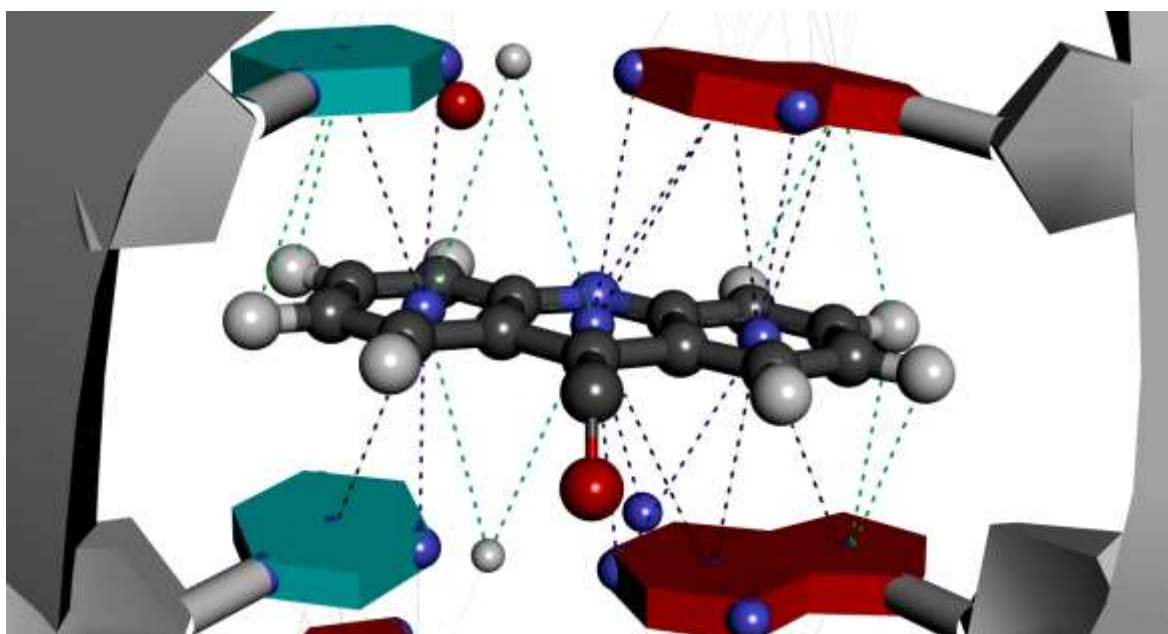
**Tabela 3.42** Interações  $\pi$  com origem na base G24, anel aromático com 6 átomos, com os átomos e centroides relacionados por bases com o respectivo comprimento **d** da interação e ângulo  $\alpha$ 

Centroide de origem da interação: G24-B								
Interage com	Átomo	Tipo de interação	1G3X DNA – LI		1VZK DNA – LS		1BNA DNA – SL	
			<b>d</b> (Å)	$\alpha$ (°)	<b>d</b> (Å)	$\alpha$ (°)	<b>d</b> (Å)	$\alpha$ (°)
C23	O2	$O \cdots \pi$			4,05	27,62		
	N3	$N \cdots \pi$	3,84	18,61	3,85	27,77	3,68	20,80
G2	N1	$N \cdots \pi$	3,66	25,87	3,89	26,04	4,07	29,80
	H21	$N - H \cdots \pi$			3,45	23,90		

Em função dos resultados do mapeamento vetorial efetuado nas estruturas de DNA analisadas, constata-se que as interações  $\pi$  sofrem influência direta tanto em função da presença de um ligante, bem como a posição em que estes se encontram. Esta afirmação é evidenciada nas Tabelas 3.7 a 3.42.

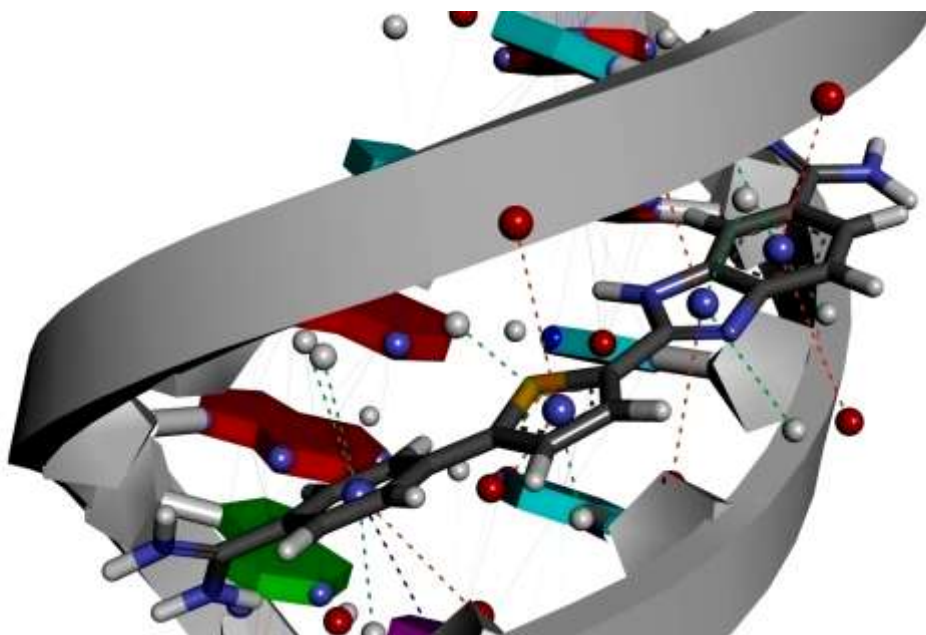
### 3.1.3 INTERAÇÕES $\pi$ ESPECÍFICAS NOS LIGANTES

As informações contidas nas Tabelas 3.7 a 3.42 permitem evidenciar que as sete interações  $\pi \cdots \pi$  existentes entre o DNA e o ligante intercalado ocorrem devido ao fato deste posicionar-se na estrutura de forma semelhante a um par de nucleotídeos, aumentando o comprimento da dupla hélice. Assim sendo, existe para esta estrutura uma reciprocidade entre as interações  $\pi$  que possuem origem tanto nos centroides do DNA quanto nos centroides do ligante (Figura 3.4).



**Figura 3.4** Representação gráfica das interações  $\pi$  entre as bases e o ligante de uma estrutura de DNA – LI, código PDB 1G3X. Estrutura gerada visualizada com o software Discovery Studio 3.1<sup>TM</sup>.

Para a estrutura de DNA com o ligante no sulco, as interações  $\pi$  existentes entre o ligante e o DNA possuem origem exclusivamente nos centroides do ligante, conforme constatado nos resultados dos cálculos efetuados e dispostos nas Tabelas 3.7 a 3.42 e que pode ser visualizado na Figura 3.5.



**Figura 3.5** Representação gráfica das interações  $\pi$  entre o ligante e as bases de uma estrutura de DNA – LS, código PDB 1VZK. Estrutura gerada visualizada com o software Discovery Studio 3.1™.

As interações  $\pi$  existentes entre as estruturas de DNA e os ligantes, em função das análises efetuadas, evidenciam o modo como a presença e a localização de um ligante altera a conformação espacial e conseqüentemente as propriedades da molécula.

### 3.2 DEFORMAÇÃO DO DNA EM FUNÇÃO DA PRESENÇA DE LIGANTES

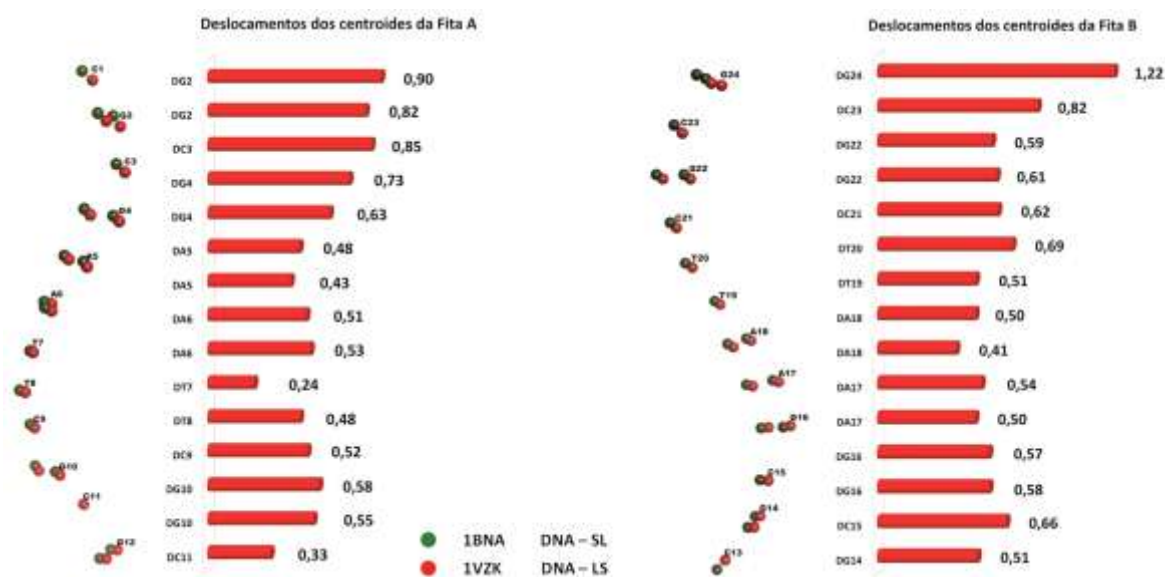
Uma vez constatado que a presença de ligantes altera a geometria espacial do DNA, faz-se necessário quantificar essa alteração para que seja entendido o quanto a posição que o ligante ocupa na estrutura afeta a sua conformação. Com base nas coordenadas dos centroides para as três estruturas obtidas pelo mapeamento vetorial (Anexos 3; 5 e 7) e utilizando como referência a estrutura de DNA sem ligante, é possível o cálculo da deformação que o ligante gera para cada posição ocupada. Com base nesse raciocínio, a deformação é mensurada em função da distância que o centroide de um mesmo anel é deslocado quando existe a presença de um ligante. Os cálculos efetuados encontram-se no Anexo 3.13.

A Tabela 3.43 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.

**Tabela 3.43** 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

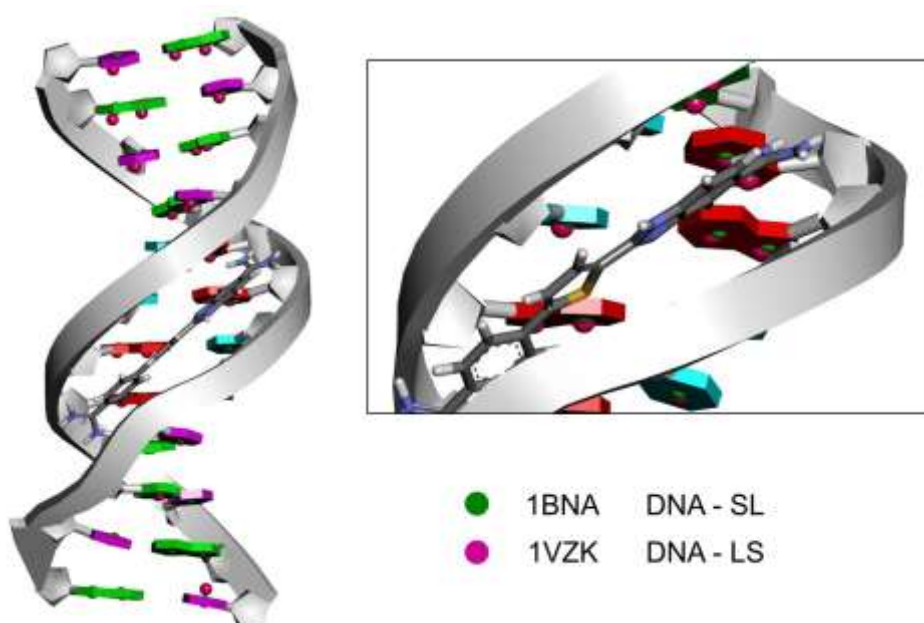
CT	Base	d	CT	Base	d
1	DC1	1,00	19	DC13	1,26
2	DG2	0,90	20	DG14	0,61
3	DG2	0,82	21	DG14	0,51
4	DC3	0,85	22	DC15	0,66
5	DG4	0,73	23	DG16	0,58
6	DG4	0,63	24	DG16	0,57
7	DA5	0,48	25	DA17	0,50
8	DA5	0,43	26	DA17	0,54
9	DA6	0,51	27	DA18	0,41
10	DA6	0,53	28	DA18	0,50
11	DT7	0,24	29	DT19	0,51
12	DT8	0,48	30	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

Os valores obtidos na Tabela 3.43 são utilizados para gerar um gráfico que ao ser colocado ao lado da imagem obtida em software gráfico dos centroides de ambas as estruturas de DNA, permitem a visualização da deformação que foi matematicamente mensurada (Figura 3.6). É evidenciado que a deformação imposta pelo ligante no sulco não apresenta algum tipo de simetria, o que é observado não apenas pelos valores do deslocamento, mas também pelas curvas do gráfico.



**Figura 3.6** Visualização gráfica da deformação calculada em função da presença de ligante no sulco da estrutura DNA – LS, tendo como base para comparação a estrutura DNA – SL.

Para um melhor entendimento, os centroides calculados de ambas as estruturas e também o ligante foram sobrepostos à visualização gráfica do DNA sem ligante (Figura 3.7). Dessa forma é possível a compreensão da deformação que o ligante exerce sobre o DNA, que é constatado pelos centroides em novas posições.



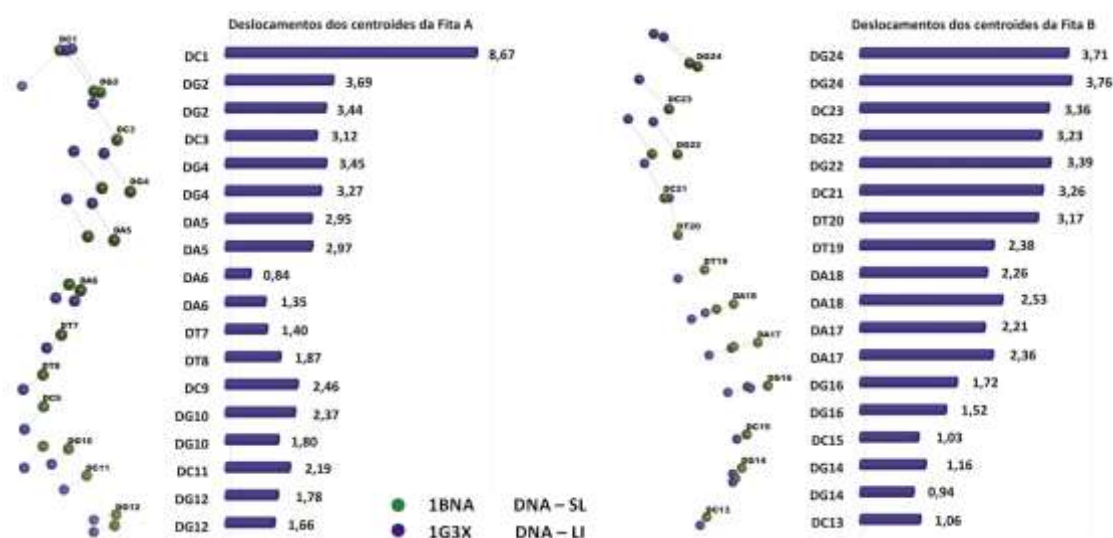
**Figura 3.7** Visualização gráfica do deslocamento dos centroides em função da presença de ligante no sulco da estrutura DNA – LS, tendo como base para comparação a estrutura DNA – SL, utilizando-se o software Discovery Studio 3.1™ para visualização. Em destaque ao lado da estrutura é visto o ligante.

Seguindo o raciocínio anterior, na Tabela 3.44 são dispostos os valores dos deslocamentos dos centroides das bases da estrutura de DNA com o ligante intercalado, tendo como referência a estrutura sem ligante.

**Tabela 3.44** 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

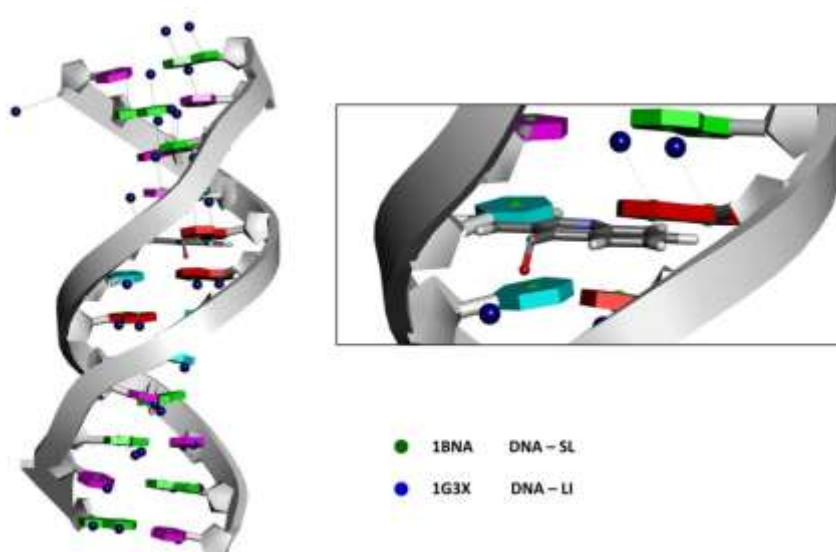
CT	Base	d	CT	Base	d
1	DC1	8,67	19	DC13	1,06
2	DG2	3,69	20	DG14	0,94
3	DG2	3,44	21	DG14	1,16
4	DC3	3,12	22	DC15	1,03
5	DG4	3,45	23	DG16	1,52
6	DG4	3,27	24	DG16	1,72
7	DA5	2,95	25	DA17	2,36
8	DA5	2,97	26	DA17	2,21
9	DA6	0,84	27	DA18	2,53
10	DA6	1,35	28	DA18	2,26
11	DT7	1,40	29	DT19	2,38
12	DT8	1,87	30	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

Comparando os valores da tabela acima com os da Tabela 3.43, evidencia-se que os valores dos deslocamentos **d** da Tabela 3.44 além de serem maiores, apresentam uma média majorada em 4,9 vezes e o desvio padrão dos deslocamentos indica uma não uniformidade na deformação 5,2 vezes maior. A interpretação gráfica dos resultados da Tabela 3.44 é representada na Figura 3.8.



**Figura 3.8** Visualização gráfica da deformação calculada em função da presença de ligante intercalado estrutura DNA – LI, tendo como base para comparação a estrutura DNA – SL.

A deformação acentuada da estrutura do DNA na região do ligante intercalado pode ser observada na Figura 3.9, onde o ligante e parte da estrutura do DNA foram sobrepostas à estrutura sem ligante, utilizada como padrão. Essa pose permite a compreensão de que o maior grau de deformação é justificado pelo fato do ligante ao assumir a posição intercalada na estrutura de DNA ocupa a posição original das bases do DNA sem ligante.

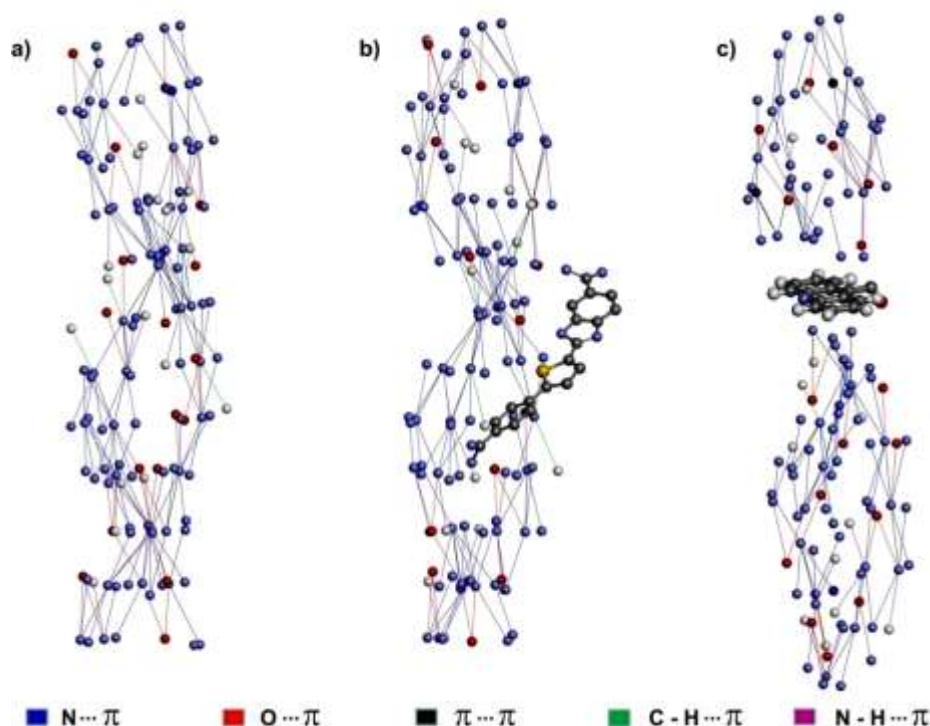


**Figura 3.9** Visualização gráfica do deslocamento dos centroides em função da presença de ligante intercalado na estrutura DNA – LI, tendo como base para comparação a estrutura DNA – SL, utilizando-se o software Discovery Studio 3.1<sup>TM</sup> para visualização. Em destaque ao lado da estrutura é visto o ligante.

Observa-se na Figura 3.8 que o DNA que possui o ligante intercalado apresenta uma deformação das duas fitas de sua estrutura com uma assimetria mais acentuada do que aquele que possui o ligante no sulco.

### 3.3 CAMINHOS ELETROSTÁTICOS DAS INTERAÇÕES $\pi$

Segundo alguns pesquisadores (Wintjens, et al., 2000; Cauët et al., 2005; Hunter e Sanders, 1990), as interações  $\pi$  contribuem para a conformação e estabilidade das proteínas, o que ocorre por meio da relação doação e recepção de cargas eletrostáticas (Kiviniemi, 2001; Hunter, et al., 2001). Com base no resultado do mapeamento vetorial para cada estrutura analisada, quando as interações  $\pi$  são identificadas, é possível elaborar um modelo no qual pode-se visualizar as relações de transferência de cargas eletrostáticas em todos os sistemas  $\pi$ . A Figura 3.10 demonstra como a presença de ligantes interfere nos sistemas  $\pi$ , alterando a configuração dos mesmos. As interações entre os ligantes e estruturas de DNA foram suprimidas.

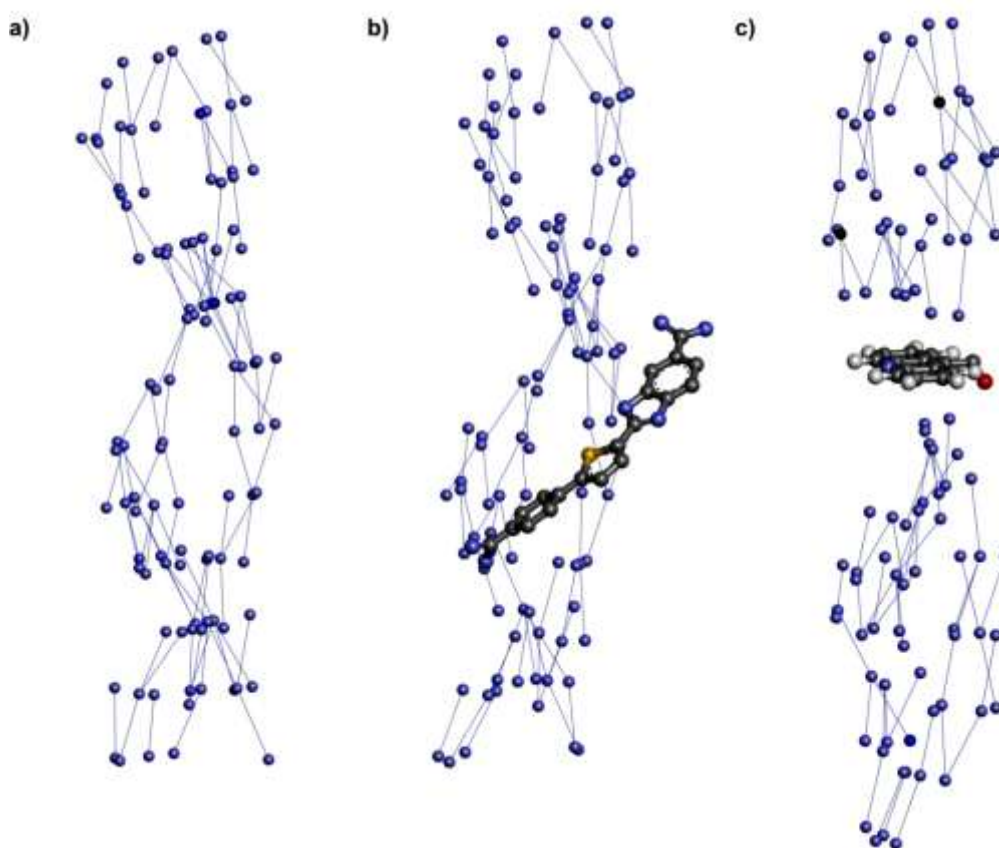


**Figura 3.10** Diagrama de transferência de cargas eletrostáticas nos sistemas  $\pi$  nas estruturas de DNA. **a)** DNA – SL; **b)** DNA – LS e **c)** DNA – LI. Utilizando o software Discovery Studio 3.1<sup>TM</sup> para visualização.



Os deslocamentos dos centroides apontados nas Tabelas 3.42 e 3.43 indicam as interações  $\pi$  que ocorrem nas estruturas em função da presença dos ligantes e podem permanecer com alteração de posição espacial ou ainda dar origem a novas interações, conforme disposto nas Tabelas 3.2 e 3.4. A visualização dessas variações por tipo de interação, excluídas as dos ligantes, conforme relacionadas nas Tabelas 3.6 a 3.42 são objeto das Figuras 3.11 a 3.20.

Para comparação das variações das interações  $\pi$  por tipo, as existentes para os dois tipos de ligantes foram sobrepostas às interações existentes no DNA sem ligante.

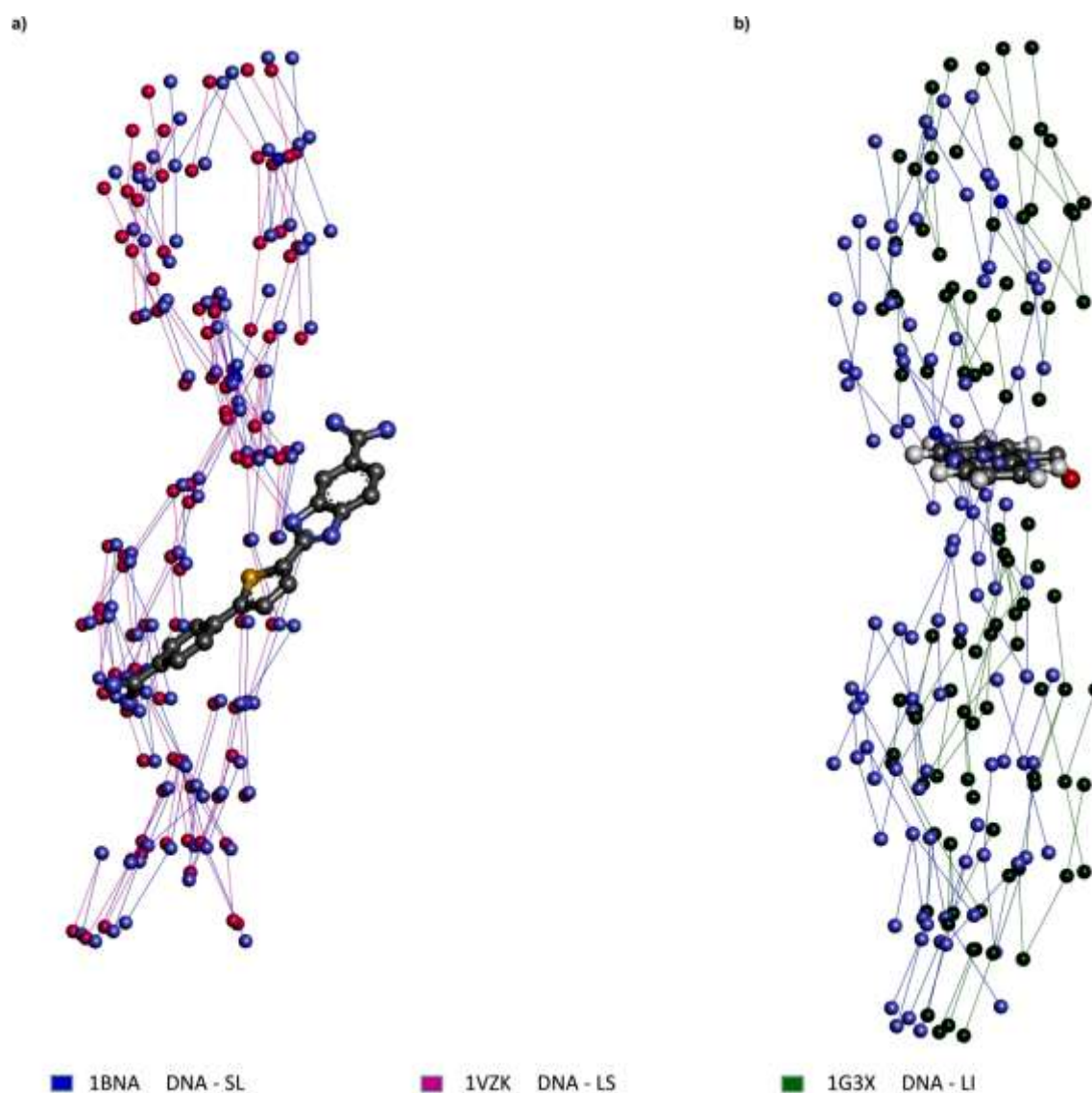


**Figura 3.11** Diagrama de transferência de cargas eletrostáticas nas interações  $N\cdots\pi$  nas estruturas de DNA. A presença dos ligantes tem como objetivo ilustrar a referência posicional destes. **a)** DNA – SL; **b)** DNA – LS e **c)** DNA – LI. Utilizado o software Discovery Studio 3.1<sup>TM</sup> para visualização.

A Tabela 3.45 apresenta o total de interações  $N\cdots\pi$  presentes nas três estruturas estudadas, não sendo computadas as interações envolvendo os ligantes. Estas interações estão representadas nas Figuras 3.11 e 3.12.

**Tabela 3.45** Comparativo das interações  $N^{\pi}\pi$  identificadas por estrutura de DNA, não computadas as interações presentes nos ligantes

Tipo de interação $\pi$	1BNA DNA – SL	1G3X DNA – LI	1VZK DNA – LS
$N^{\pi}\pi$	88	75	81

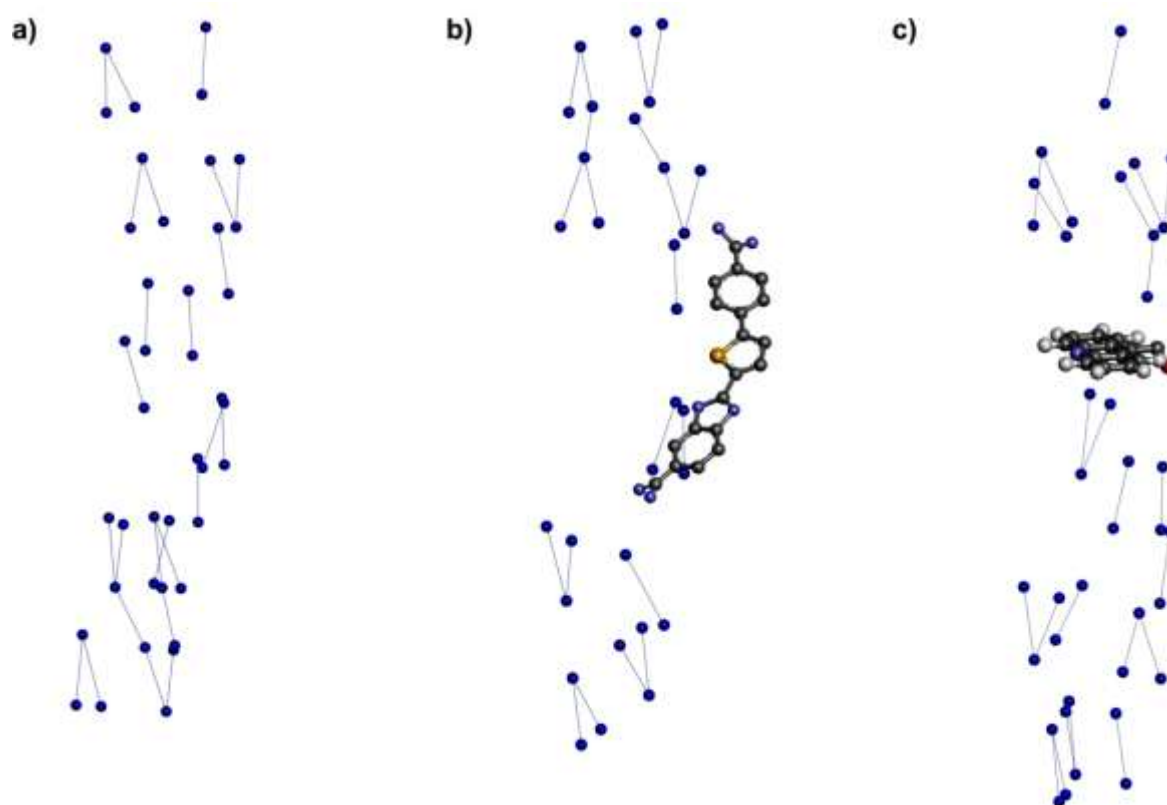


**Figura 3.12** Sobreposição das interações  $N^{\pi}\pi$  das estruturas de DNA com ligantes sobrepostas à estrutura DNA – SL. A presença dos ligantes tem como objetivo ilustrar a sua posição. **a)** DNA – LS. e **b)** DNA – LI. Utilizado o software Discovery Studio 3.1<sup>TM</sup> para visualização.

Estão contidas na Tabela 3.46 o total de interações  $O^{\delta-}\pi$  presentes nas três estruturas estudadas. Estas interações estão representadas nas Figuras 3.13 e 3.14.

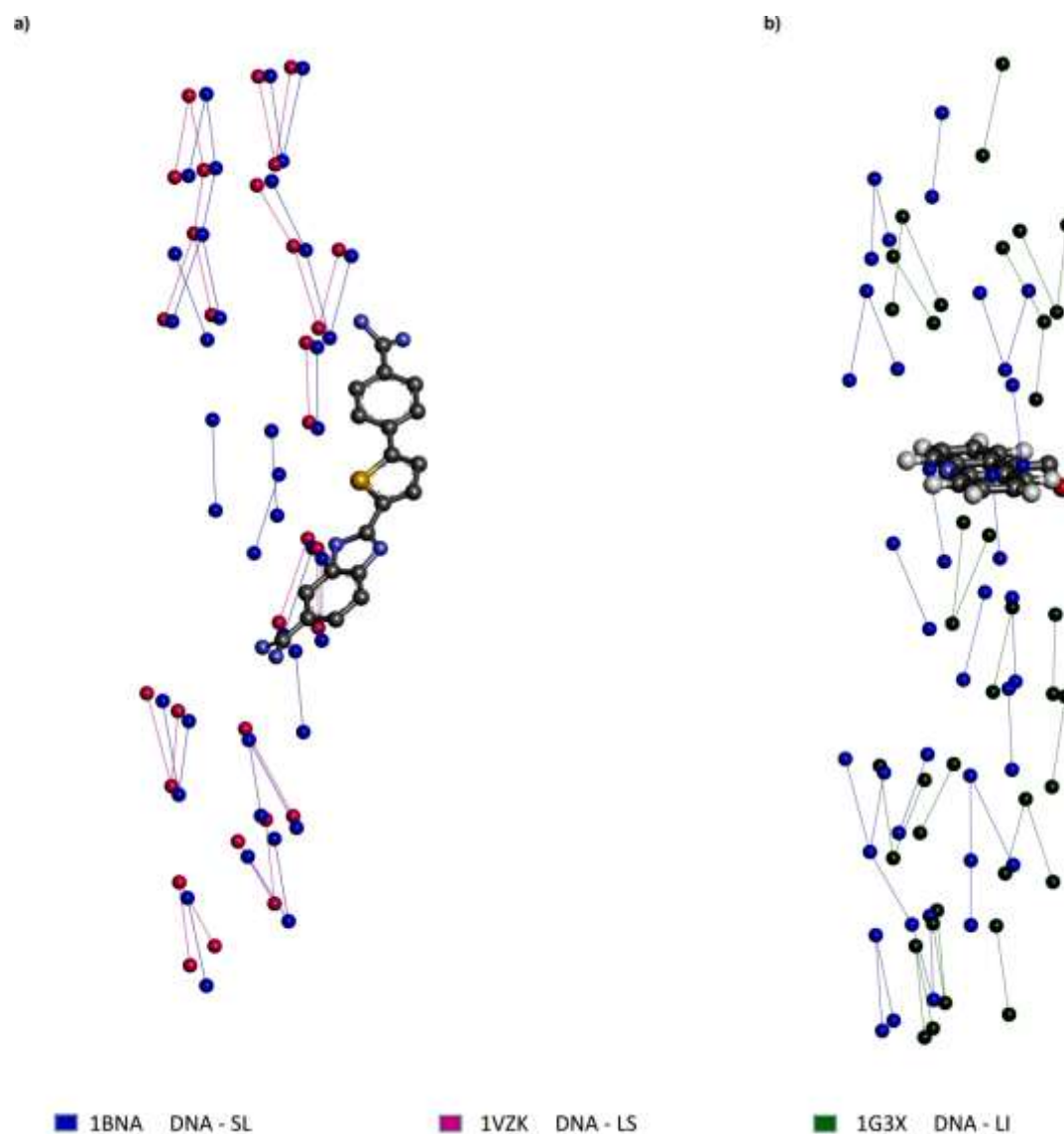
**Tabela 3.46** Comparativo das interações  $O^{\delta-}\pi$  identificadas por estrutura de DNA, não computadas as interações presentes nos ligantes.

Tipo de interação $\pi$	1BNA DNA – SL	1G3X DNA – LI	1VZK DNA – LS
$O^{\delta-}\pi$	25	23	24



**Figura 3.13** Diagrama de transferência de cargas eletrostáticas nas interações  $O^{\delta-}\pi$  nas estruturas de DNA. A presença dos ligantes tem como objetivo ilustrar a referência posicional destes. **a)** DNA – SL; **b)** DNA – LS e **c)** DNA – LI. Utilizado o software Discovery Studio 3.1<sup>TM</sup> para visualização.

Pode-se visualizar na Figura 3.13 que a distribuição uniforme das interações  $O\cdots\pi$  ao longo da estrutura do DNA sem ligante é alterada nas demais estruturas nas regiões adjacentes aos ligantes.

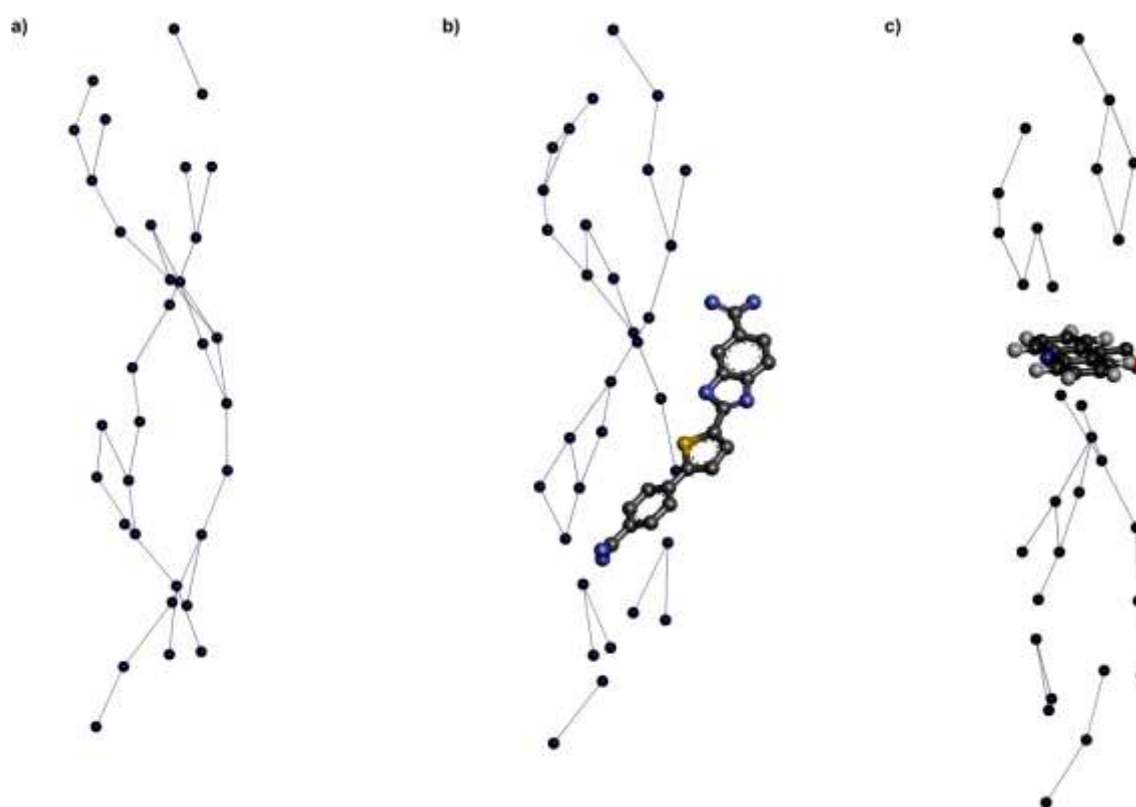


**Figura 3.14** Sobreposição das interações  $O\cdots\pi$  das estruturas de DNA com ligantes sobrepostas à estrutura DNA - SL. A presença dos ligantes tem como objetivo ilustrar a sua posição. a) DNA - LS. e b) DNA - LI. Utilizado o software Discovery Studio 3.1<sup>TM</sup> para visualização.

Na Tabela 3.47 estão relacionadas o total de interações  $\pi\cdots\pi$  presentes nas três estruturas estudadas. Estas interações estão representadas nas Figuras 3.15 e 3.16.

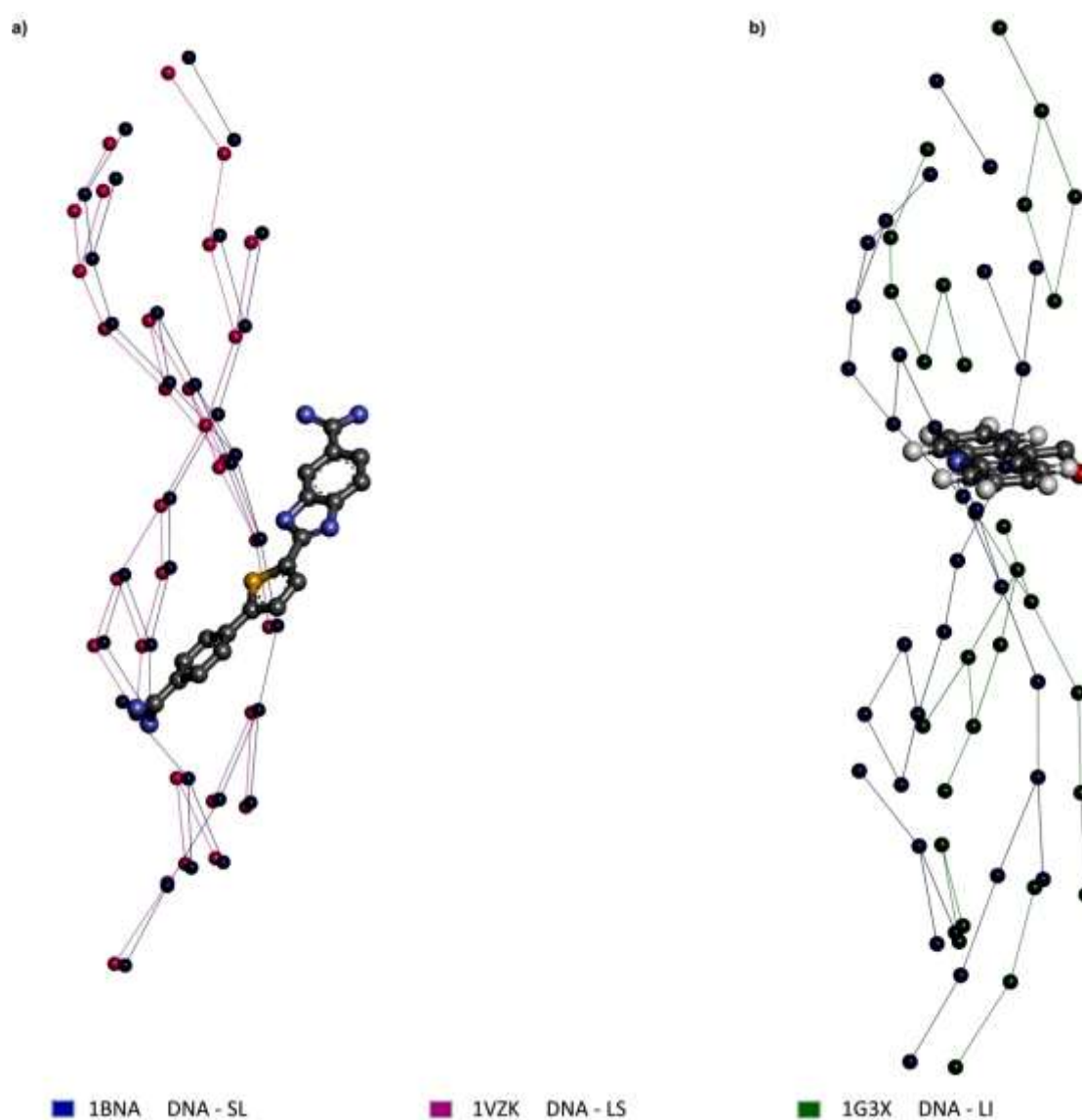
**Tabela 3.47** Comparativo das interações  $\pi\cdots\pi$  identificadas por estrutura de DNA, não computadas as interações presentes nos ligantes.

Tipo de interação $\pi$	1BNA DNA – SL	1G3X DNA – LI	1VZK DNA – LS
$\pi\cdots\pi$	32	25	31



**Figura 3.15** Diagrama de transferência de cargas eletrostáticas nas interações  $\pi\cdots\pi$  nas estruturas de DNA. A presença dos ligantes tem como objetivo ilustrar a referência posicional destes. **a)** DNA – SL; **b)** DNA – LS e **c)** DNA – LI. Utilizado o software Discovery Studio 3.1<sup>TM</sup> para visualização.

Embora exista uma ausência de interações  $\pi\cdots\pi$  na região dos ligantes, a forma geométrica das interações ao longo das três estruturas apresenta linhas comuns (Figura 3.15).

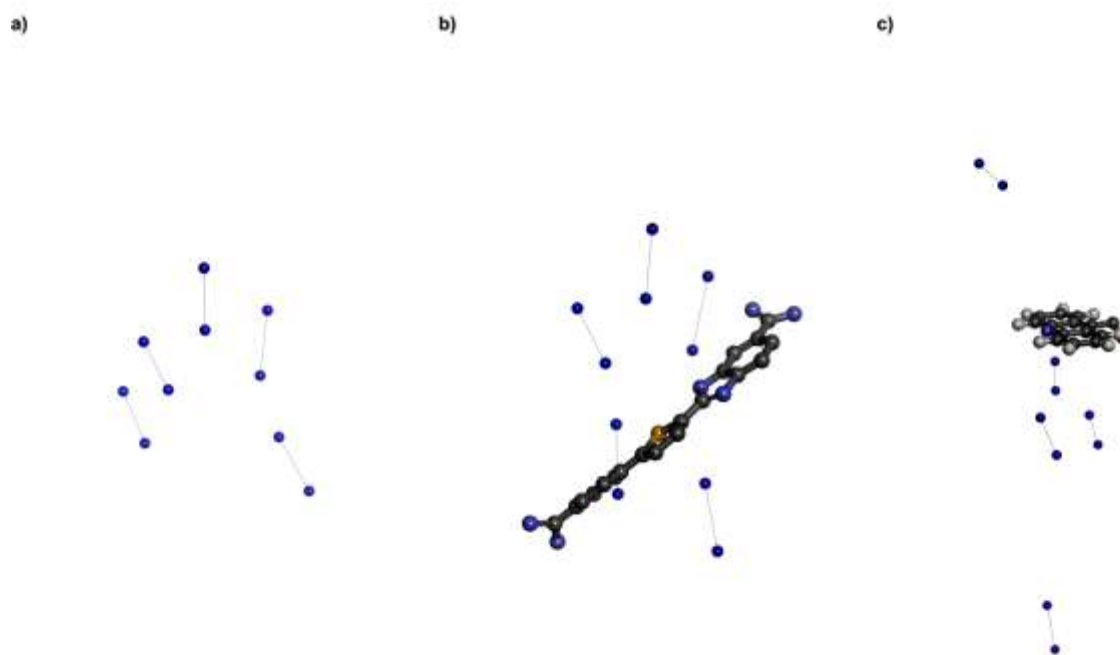


**Figura 3.16** Sobreposição das interações  $\pi\cdots\pi$  das estruturas de DNA com ligantes sobrepostas à estrutura DNA - SL. A presença dos ligantes tem como objetivo ilustrar a sua posição. **a)** DNA - LS. e **b)** DNA - LI. Utilizado o software Discovery Studio 3.1<sup>TM</sup> para visualização.

Relacionam-se na Tabela 3.48 o total de interações C-H $\cdots\pi$  presentes nas três estruturas estudadas. Estas interações estão representadas nas Figuras 3.17 e 3.18.

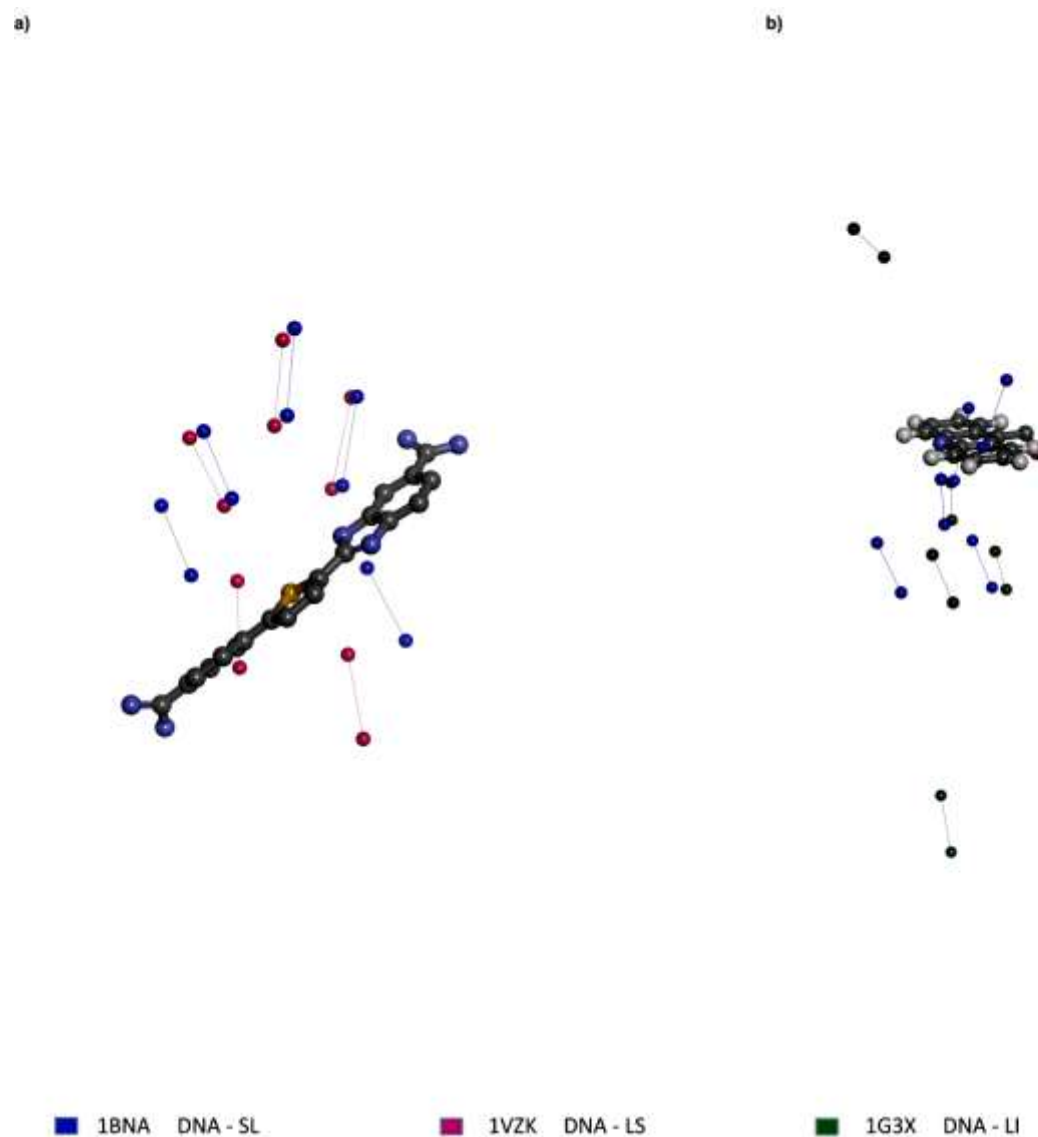
**Tabela 3.48** Comparativo das interações C-H $\cdots\pi$  identificadas por estrutura de DNA, não computadas as interações presentes nos ligantes.

Tipo de interação $\pi$	1BNA DNA – SL	1G3X DNA – LI	1VZK DNA – LS
C - H $\cdots\pi$	5	5	5



**Figura 3.17** Diagrama de transferência de cargas eletrostáticas nas interações C-H $\cdots\pi$  nas estruturas de DNA. A presença dos ligantes tem como objetivo ilustrar a referência posicional destes. **a)** DNA – SL; **b)** DNA – LS e **c)** DNA – LI. Utilizado o software Discovery Studio 3.1<sup>TM</sup> para visualização.

Pode-se constatar na Tabela 3.48 que existe quantitativamente a redução de uma interação C-H $\cdots$  $\pi$  para o DNA com o ligante no sulco, a disposição dessas interações não apresenta um padrão de localização comum às três estruturas (Figura 3.17).



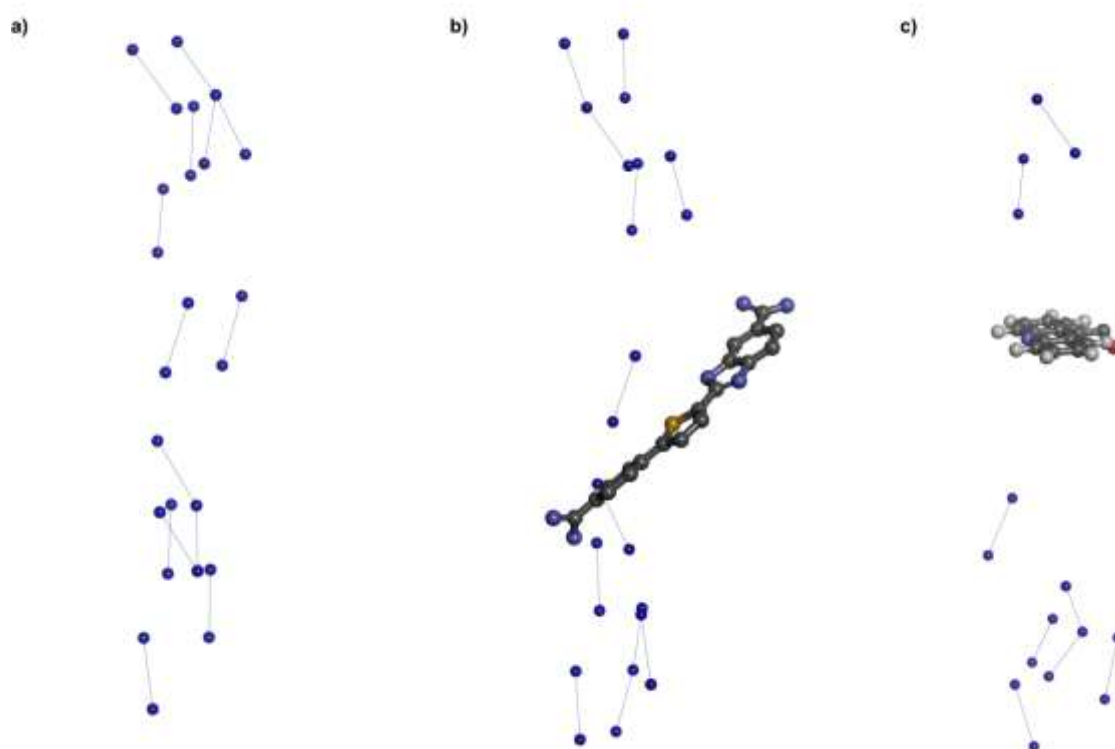
**Figura 3.18** Sobreposição das interações C-H $\cdots$  $\pi$  das estruturas de DNA com ligantes sobrepostas à estrutura DNA – SL. A presença dos ligantes tem como objetivo ilustrar a sua posição. **a)** DNA – LS. e **b)** DNA – LI. Utilizado o software Discovery Studio 3.1<sup>TM</sup> para visualização.



Na Tabela 3.49 estão as interações N-H $\cdots\pi$  presentes nas três estruturas de DNA. Estas interações estão representadas nas Figuras 3.19 e 3.20.

**Tabela 3.49** Comparativo das interações N-H $\cdots\pi$  identificadas por estrutura de DNA, não computadas as interações presentes nos ligantes.

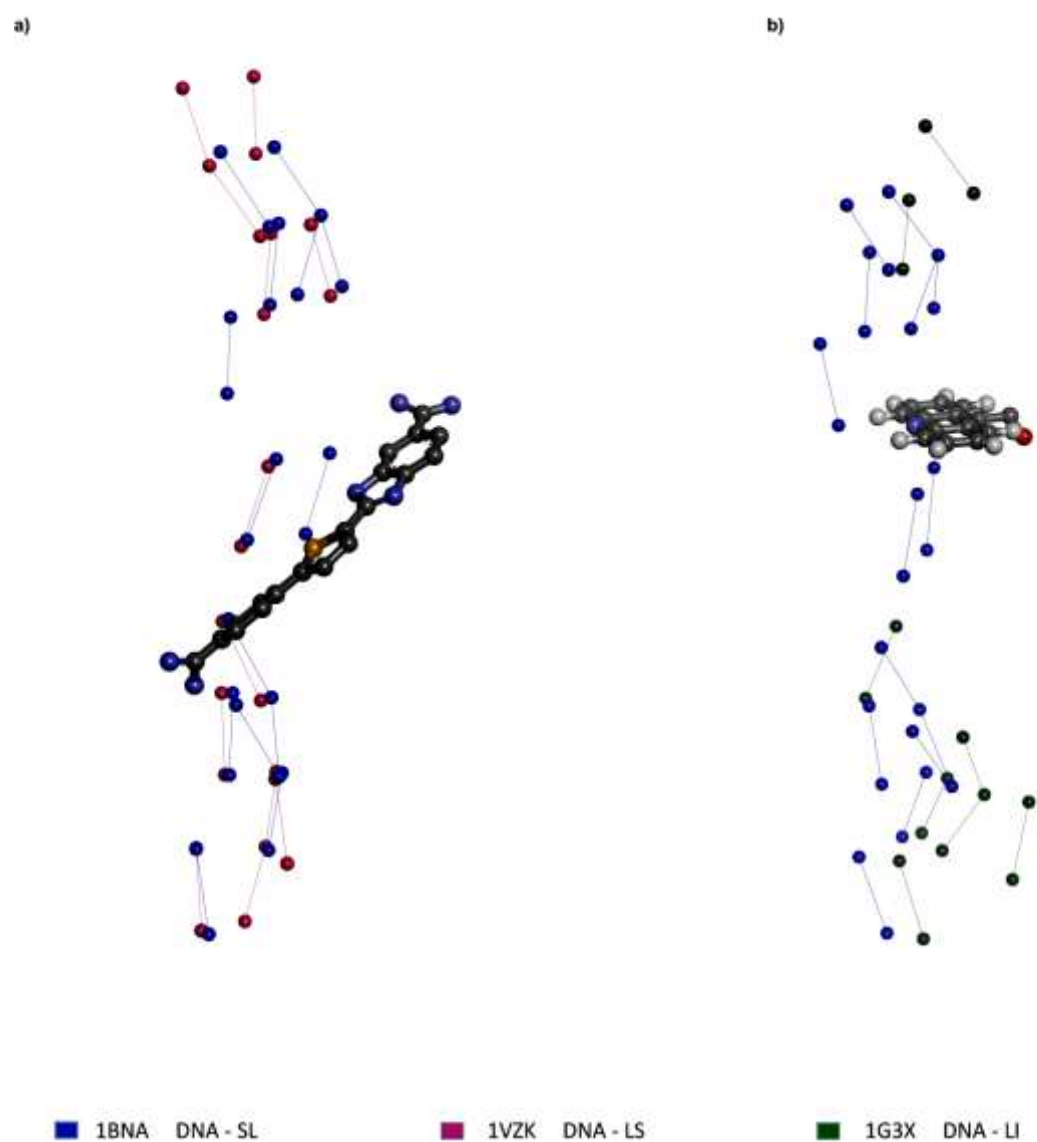
Tipo de interação $\pi$	1BNA DNA – SL	1G3X DNA – LI	1VZK DNA – LS
N - H $\cdots\pi$	14	9	13



**Figura 3.19** Diagrama de transferência de cargas eletrostáticas nas interações N-H $\cdots\pi$  nas estruturas de DNA. A presença dos ligantes tem como objetivo ilustrar a referência posicional destes. **a)** DNA – SL; **b)** DNA – LS e **c)** DNA – LI. Utilizado o software Discovery Studio 3.1<sup>TM</sup> para visualização.

A deformação causada pela presença do ligante intercalado na estrutura do DNA acarreta numa redução quantitativa das interações superior a 30% em relação às

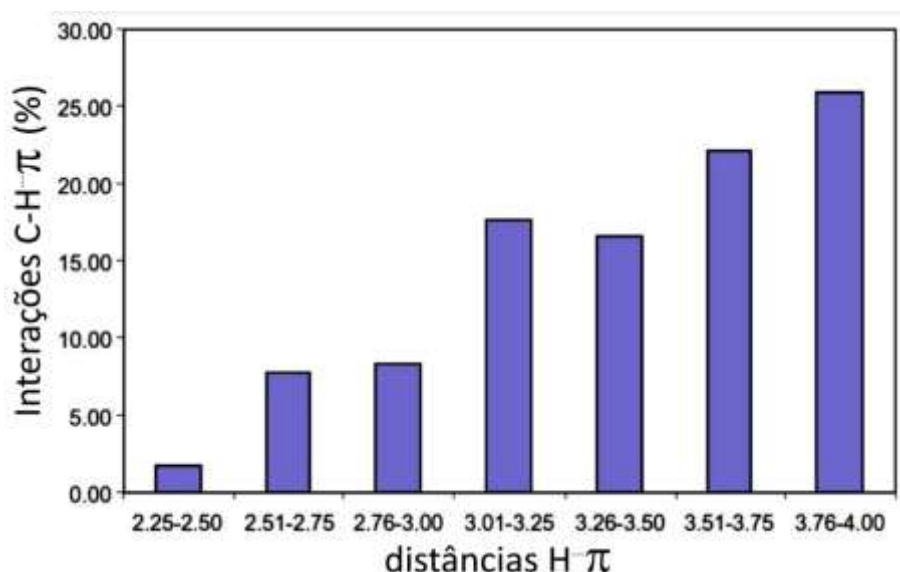
demais estruturas de DNA, enquanto que a sem ligante e a com ligante no sulco apresentam padrões geométricos semelhantes (Figuras 3.19 e 3.20).



**Figura 3.20** Sobreposição das interações N-H... $\pi$  das estruturas de DNA com ligantes sobrepostas à estrutura DNA - SL. A presença dos ligantes tem como objetivo ilustrar a sua posição. **a)** DNA - LS. e **b)** DNA - LI. Utilizado o software Discovery Studio 3.1<sup>TM</sup> para visualização.

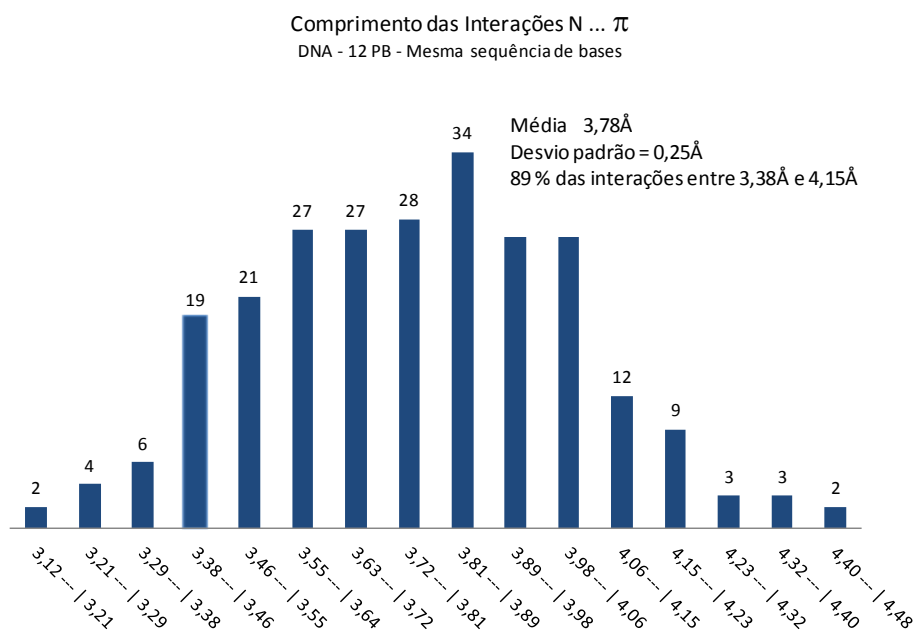
### 3.4 ESTATÍSTICA DOS PARÂMETROS GEOMÉTRICOS DAS INTERAÇÕES $\pi$

Ao utilizar o mapeamento vetorial para determinação das interações  $\pi$  em macromoléculas e DNA, além da identificação das interações válidas, os cálculos retornam os valores dos parâmetros geométricos das distâncias e ângulos dos vetores distâncias relativos aos vetores normais aos planos dos anéis. Estes parâmetros servem também como base para uma análise estatística que permite identificar características para cada tipo de interação. Esse tipo de análise foi utilizado inicialmente por Shanthi et al. (2010) no estudo das interações C-H $\cdots\pi$  existentes na cadeia lateral de uma amostragem de 75 proteínas, onde tabulou o comprimento das interações entre os átomos de hidrogênio e os centroides (Figura 3.21).

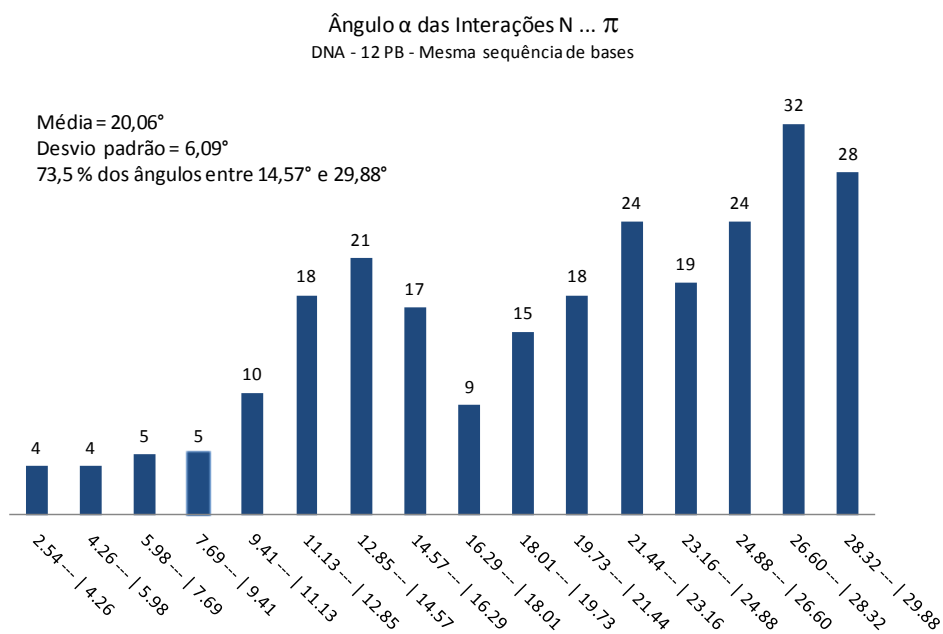


**Figura 3.21** Representação gráfica da tabulação das distâncias entre os átomos de hidrogênio e centroides nas interações C-H $\cdots\pi$  da cadeia lateral de proteínas. Extraído de Shanthi et al. (2010).

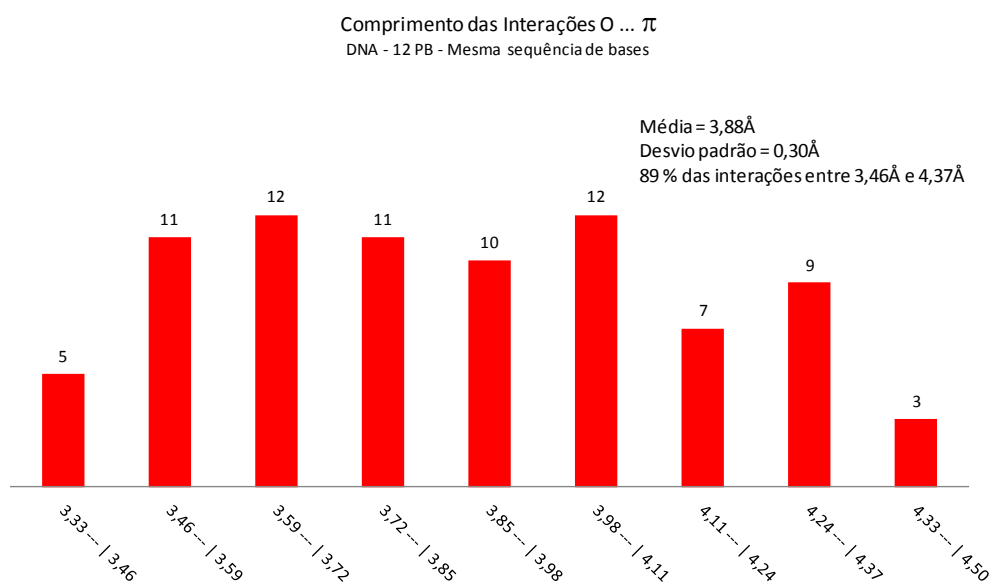
No Anexo 3.8 foi efetuada a tabulação dos dados referentes às interações  $\pi$  calculadas para as três estruturas de DNA estudadas, incluindo as que envolvem os ligantes. Os resultados incluem a média dos parâmetros geométricos, o desvio padrão da amostra e o intervalo na qual estão concentradas a maioria das medidas obtidas (Figuras 3.22 a 3.31).



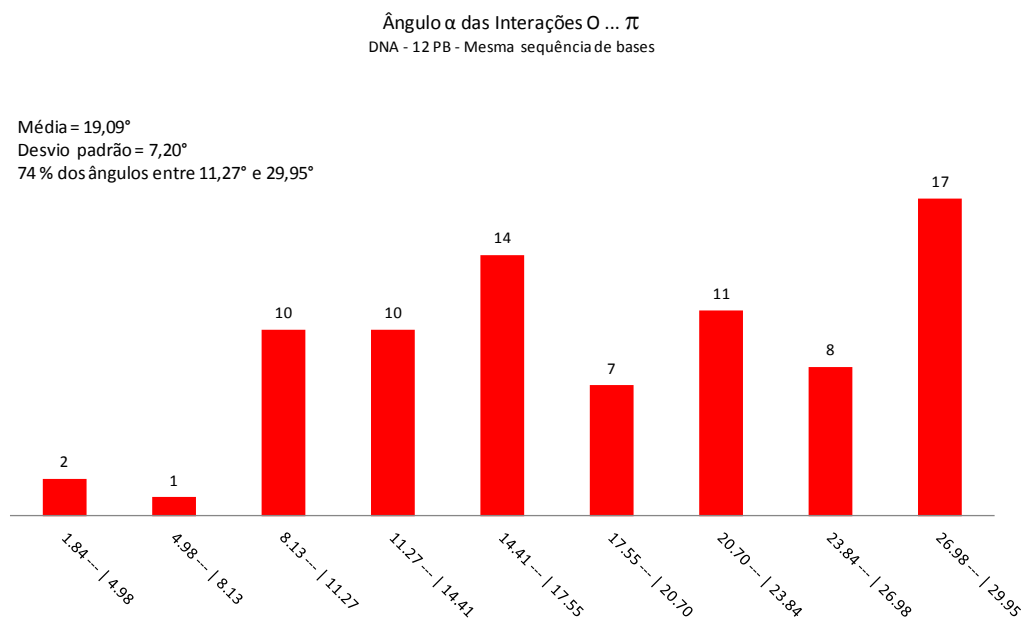
**Figura 3.22** Representação gráfica da tabulação das distâncias das interações N... $\pi$  das estruturas 1BNA (DNA – SL); 1VZK (DNA – LS) e 1G3X (DNA – LI).



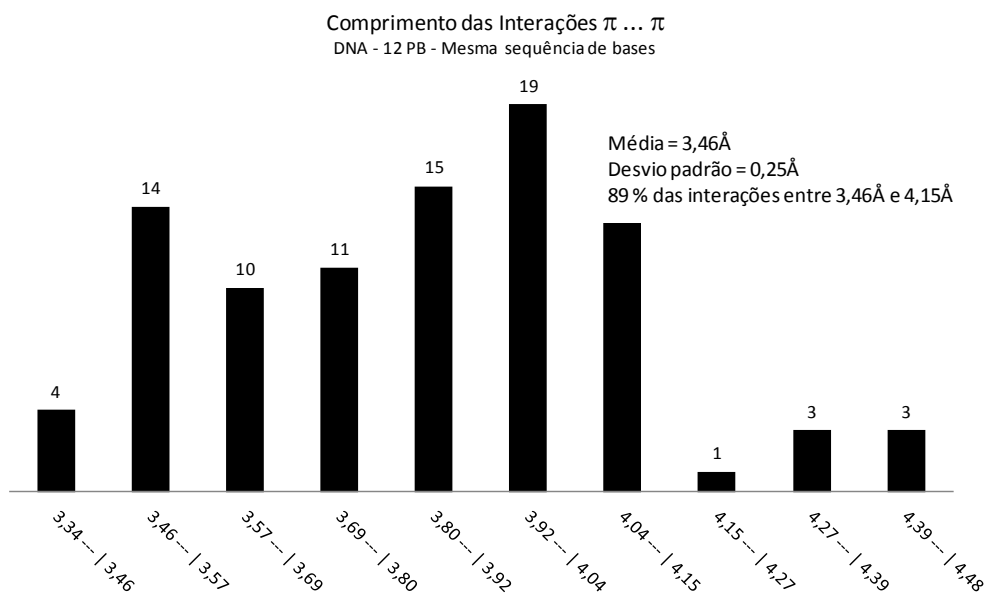
**Figura 3.23** Representação gráfica da tabulação dos ângulos entre o vetor distância e o vetor normal ao plano do anel aromático das interações N... $\pi$  das estruturas 1BNA (DNA – SL); 1VZK (DNA – LS) e 1G3X (DNA – LI).



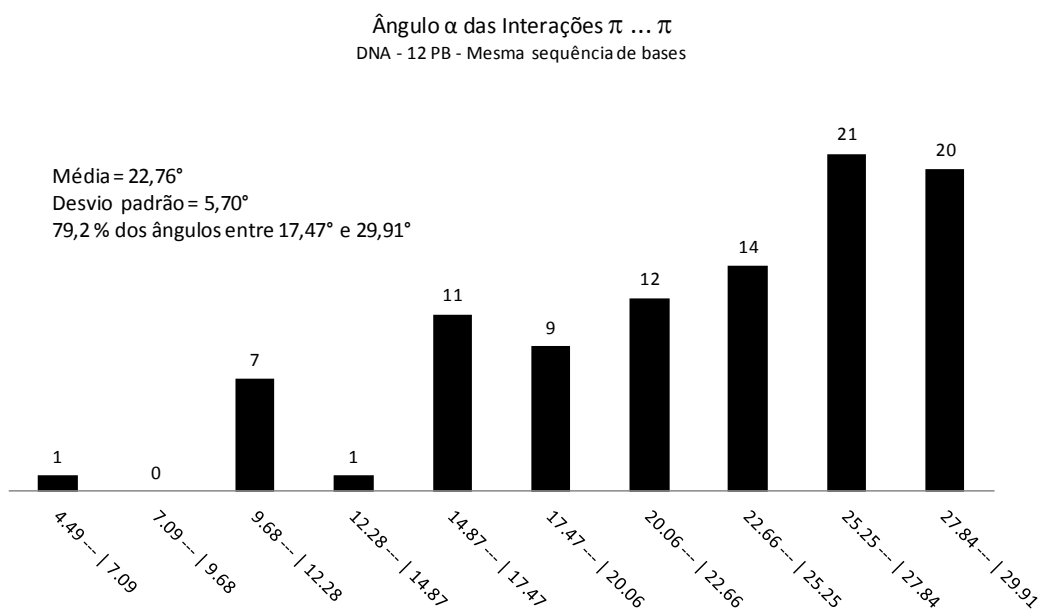
**Figura 3.24** Representação gráfica da tabulação das distâncias das interações O... $\pi$  das estruturas 1BNA (DNA – SL); 1VZK (DNA – LS) e 1G3X (DNA – LI).



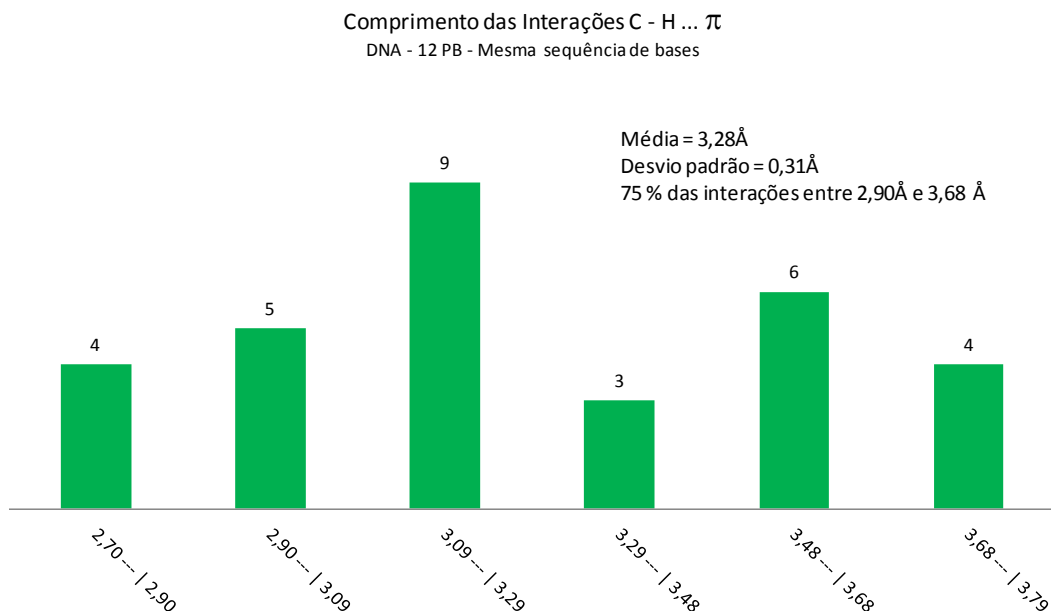
**Figura 3.25** Representação gráfica da tabulação dos ângulos entre o vetor distância e o vetor normal ao plano do anel aromático das interações O... $\pi$  das estruturas 1BNA (DNA – SL); 1VZK (DNA – LS) e 1G3X (DNA – LI).



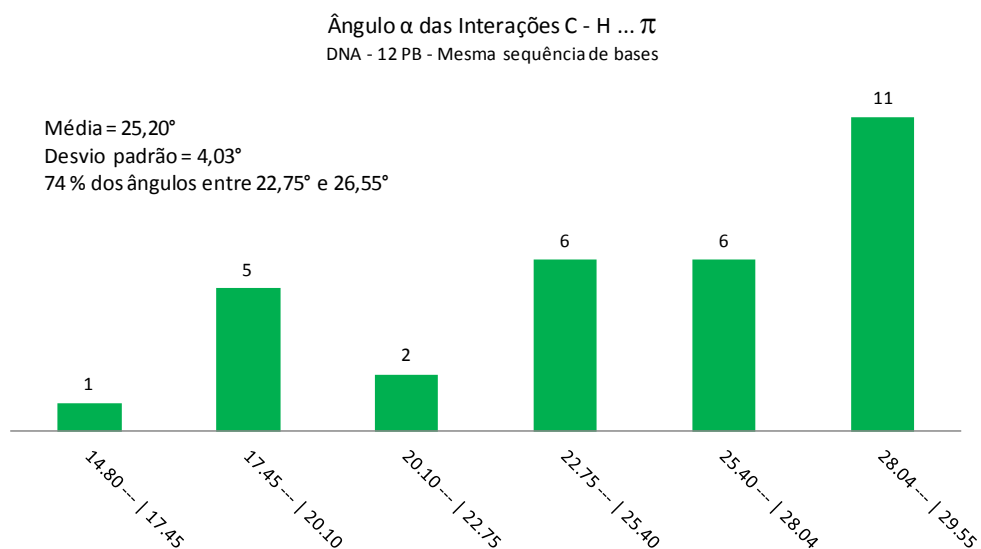
**Figura 3.26** Representação gráfica da tabulação das distâncias das interações  $\pi \dots \pi$  das estruturas 1BNA (DNA – SL); 1VZK (DNA – LS) e 1G3X (DNA – LI).



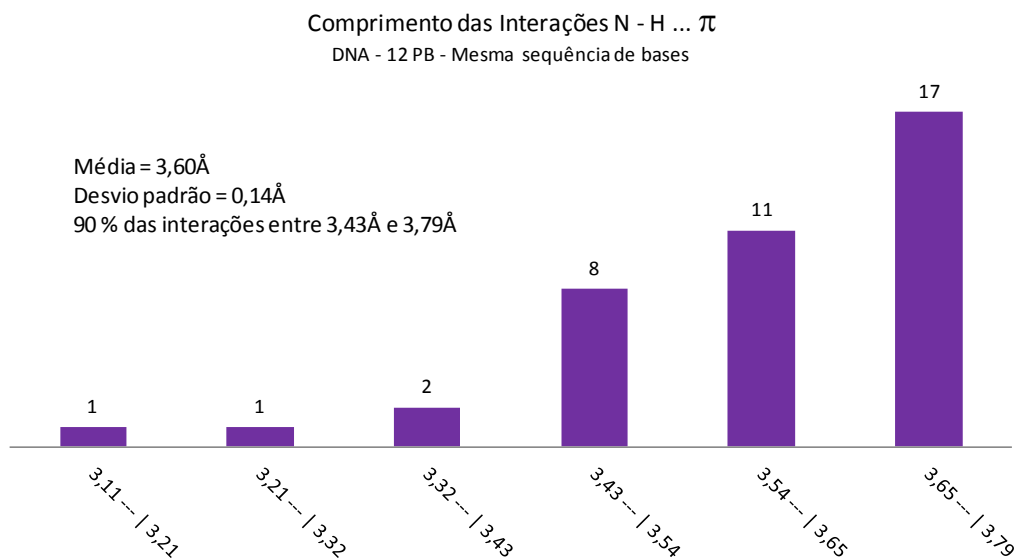
**Figura 3.27** Representação gráfica da tabulação dos ângulos entre o vetor distância e o vetor normal ao plano do anel aromático das interações  $\pi \dots \pi$  das estruturas 1BNA (DNA – SL); 1VZK (DNA – LS) e 1G3X (DNA – LI).



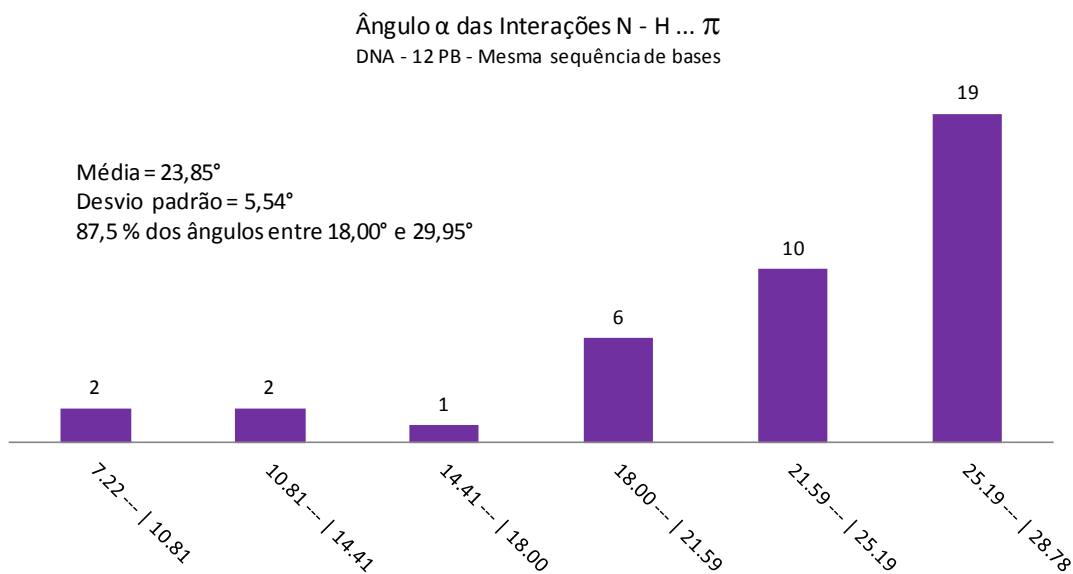
**Figura 3.28** Representação gráfica da tabulação das distâncias das interações C-H... $\pi$  das estruturas 1BNA (DNA – SL); 1VZK (DNA – LS) e 1G3X (DNA – LI).



**Figura 3.29** Representação gráfica da tabulação dos ângulos entre o vetor distância e o vetor normal ao plano do anel aromático das interações C-H... $\pi$  das estruturas 1BNA (DNA – SL); 1VZK (DNA – LS) e 1G3X (DNA – LI).



**Figura 3.30** Representação gráfica da tabulação das distâncias das interações N-H... $\pi$  das estruturas 1BNA (DNA – SL); 1VZK (DNA – LS) e 1G3X (DNA – LI).



**Figura 3.31** Representação gráfica da tabulação dos ângulos entre o vetor distância e o vetor normal ao plano do anel aromático das interações N-H... $\pi$  das estruturas 1BNA (DNA – SL); 1VZK (DNA – LS) e 1G3X (DNA – LI).



Uma vez sendo conhecidos os valores calculados das distâncias das interações  $\pi$  e os ângulos  $\alpha$  existentes entre o vetor distância e o vetor normal ao plano do anel, é possível efetuar uma comparação entre os valores teóricos e os obtidos pelo mapeamento vetorial para os parâmetros geométricos estudados. A Tabela 3.50 apresenta os dados comparativos.

**Tabela 3.50** Comparativo entre os valores teóricos e os obtidos pelo mapeamento vetorial para os parâmetros geométricos para as interações  $\pi$ , por tipo das estruturas de DNA estudadas

Tipo de Interação	Parâmetros Teóricos (Tabela 2.2)		Valores obtidos pelo mapeamento vetorial	
	Módulo do vetor distância	Ângulo entre o vetor normal e o vetor distância	Módulo do vetor distância	Ângulo entre o vetor normal e o vetor distância
$\pi \cdots \pi$	$3,0 \leq d \leq 4,5 \text{ \AA}$	$\alpha \text{ e } \beta \leq 30^\circ$	$3,34 \leq d \leq 4,48 \text{ \AA}$	$4,49^\circ \leq \alpha \text{ e } \beta \leq 29,91^\circ$
C-H $\cdots\pi$	$2,5 \leq d \leq 3,8 \text{ \AA}$	$\alpha \leq 30^\circ$	$2,70 \leq d \leq 3,79 \text{ \AA}$	$14,80^\circ \leq \alpha \leq 29,55^\circ$
N-H $\cdots\pi$	$2,5 \leq d \leq 3,8 \text{ \AA}$	$\alpha \leq 30^\circ$	$3,11 \leq d \leq 3,79 \text{ \AA}$	$7,22^\circ \leq \alpha \leq 28,78^\circ$
O-H $\cdots\pi$	$2,5 \leq d \leq 3,8 \text{ \AA}$	$\alpha \leq 30^\circ$	-	-
N $\cdots\pi$	$3,2 \leq d \leq 4,5 \text{ \AA}$	$\alpha \leq 30^\circ$	$3,12 \leq d \leq 4,48 \text{ \AA}$	$2,54^\circ \leq \alpha \leq 29,88^\circ$
O $\cdots\pi$	$3,2 \leq d \leq 4,5 \text{ \AA}$	$\alpha \leq 30^\circ$	$3,33 \leq d \leq 4,50 \text{ \AA}$	$1,84^\circ \leq \alpha \leq 29,55^\circ$

### 3.5 CONCLUSÕES DO CAPÍTULO 3

O mapeamento vetorial efetuado para as três estruturas de DNA selecionadas trouxe além dos objetivos previstos inicialmente, outras aplicações que geraram informações não disponíveis na literatura. Abaixo estão relacionadas essas aplicações.

- Variação das interações  $\pi$  em função da presença de ligantes;
- Análise das variações das interações  $\pi$  por nucleotídeo para os DNA mapeados vetorialmente;
- Identificação de interações  $\pi$  específicas nos ligantes;
- Mensuração da deformação do DNA em função da presença de ligantes;

- Localização dos caminhos eletrostáticos das interações  $\pi$  para cada estrutura de DNA mapeada;
- Estatística dos parâmetros geométricos reais das interações  $\pi$  calculados vetorialmente para as estruturas de DNA estudadas.

# Capítulo 4

---

CONCLUSÕES

#### 4. CONCLUSÕES

Ao publicar seu trabalho sobre a estrutura da mioglobina, Kendrew iniciou uma nova era para o estudo das macromoléculas, sendo preconizado por ele que a estrutura tridimensional de uma macromolécula é de suma importância para a compreensão da sua função biológica. Com a evolução dos recursos computacionais e a criação do PDB, cuja função é ser um banco de dados colaborativo para depósito de estruturas tridimensionais de macromoléculas, obtidas principalmente por meio de experimentos cristalográficos, tornou-se possível a visualização em tela gráfica da conformação espacial destas estruturas. Essa relação interdisciplinar entre a biologia, ciência da computação, cristalografia e matemática tem possibilitado à química supramolecular atender aos seus objetivos quando esta necessita de ferramentas para o estudo dos arranjos moleculares nos quais a identificação de interações não covalentes é um fator preponderante. No escopo destas interações não covalentes, estão as interações  $\pi$ , uma das responsáveis pela conformação espacial de macromoléculas e em particular da dupla hélice do DNA.

Todavia, mesmo com a evolução dos softwares para a visualização gráfica das estruturas cristalinas de macromoléculas e o seu manuseio virtual em tela gráfica, a identificação das interações  $\pi$  até o momento deve ser feita de forma manual uma a uma o que exige atenção, análise visual dos parâmetros envolvidos o que demanda bastante tempo. Quando se pretende estudar os sistemas biológicos, com milhares de átomos, qualquer processo que seja automatizado, neste caso para a identificação e análise das interações  $\pi$ , vai facilitar a extração e análise dos dados procurados.

Como resposta a esse problema, este trabalho apresenta a ideia e implementação do uso de conceitos de álgebra linear para efetuar o mapeamento vetorial para identificar as interações  $\pi$  em biomoléculas. A ferramenta computacional aqui apresentada permite solucionar os inconvenientes identificados acima, servindo como ponto inicial à ação do pesquisador e direcionando a sua ação para as atividades que agregam valor à pesquisa. Dessa forma, o usuário pode converter o tempo

despendido na busca e identificação manuais das interações  $\pi$  na análise da influência das interações identificadas, de modo automático, nas funções biológicas da macromolécula.

O fluxograma criado para o processamento matemático dos arquivos referentes às estruturas cristalográficas depositadas no PDB para determinar as coordenadas dos centroides dos anéis  $\pi$  (arila) mostrou que além da relação de interações  $\pi$  existentes em estruturas de proteínas ou DNA, permite obter com exatidão outras informações relacionadas à estrutura, como por exemplo, a quantificação de uma deformação. Este tipo de informação não está disponível na literatura em função da inexistência de uma ferramenta que pudesse ser utilizada para sua obtenção. Estas análises qualitativas e quantitativas relacionadas às interações  $\pi$ , e que podem estar relacionadas com as propriedades dessas moléculas são detalhadas a seguir:

- A contabilidade das interações  $\pi$  existentes em cada estrutura e as variações absolutas e relativas existentes em função dos ligantes.
- Para cada interação identificada, foram determinados automaticamente o seu comprimento e o ângulo entre o vetor normal ao plano do anel e o vetor distância ao átomo envolvido na interação ou a outro anel e cujo módulo é o comprimento dessa interação.
- A relação desses resultados aplicados a moléculas de DNA permitiu avaliar a influência que os ligantes exercem na conformação das fitas, pela mensuração da quantidade de cada tipo de interação  $\pi$  existente, e quando estas não variam quantitativamente em função da presença de ligantes, é possível a constatação da alteração dos seus parâmetros geométricos, identificando alterações nos ângulos  $\alpha$  e comprimentos das interações.
- A utilização das coordenadas dos centroides calculados para cada anel aromático existente em cada base das estruturas estudadas como referência posicional permitiu a mensuração e visualização em tela gráfica do grau de deformação que o ligante induz na geometria da molécula.

- Deve ser destacado que pela primeira vez essa deformação foi mensurada e foi, então, possível constatar que esta não é uniforme em ambas as estruturas de DNA com ligante, ou seja, cada fita da estrutura apresenta um grau de deformação diferente.
- Outro resultado obtido foi que, uma vez quantificadas as deformações existentes, utilizando como referência as coordenadas dos centroides dos anéis aromáticos, foi possível efetuar uma análise estatística destas e assim concluir que para a estrutura com o ligante intercalado o grau de deformação é maior do que o apurado para a estrutura que possui o ligante no sulco.
- Uma vez obtida a relação de todas as interações  $\pi$  determinadas pelo mapeamento vetorial, foi possível transformá-la em um conjunto de dados estatísticos no qual as variáveis quantitativas utilizadas foram os elementos envolvidos na interação e os ângulos e comprimentos das interações para cada tipo de interação  $\pi$ .
- Dessa forma, para cada tipo de interação foi possível efetuar uma análise estatística e avaliar a distribuição das frequências de ângulos e distâncias. Em função dos histogramas obtidos e medidas de dispersão foi possível constatar que não existe um padrão comum tanto para os comprimentos das interações quanto para seus ângulos.
- Com base na amostra de dados referentes aos ângulos e distâncias das interações identificadas foi possível efetuar uma comparação com os parâmetros geométricos teóricos obtidos no trabalho de pesquisa, estando os resultados obtidos contidos nos intervalos teóricos.

Assim, os resultados obtidos no desenvolvimento deste trabalho mostram a utilidade do mapeamento vetorial para a pesquisa das interações supramoleculares que envolvem sistemas aromáticos e justificam a continuidade da pesquisa de forma a transformar o fluxograma desenvolvido em um software que possa ser utilizado como ferramenta de pesquisa. Esse software, além de automatizar o trabalho de identificação das interações  $\pi$  com exatidão, reduzindo o tempo da pesquisa e proporcionando ao pesquisador um aumento da sua produtividade, permitirá o acesso a informações que até

o presente momento não eram disponíveis. 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.

# Referências

---



Accelrys Discovery Studio Visualizer™ - Accelrys Software Inc. Discovery Studio Visualizer Client 3.1 - <http://accelrys.com/products/discovery-studio/index.html>

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.

Batista, E. D. O.: Sistemas de Informação: o uso consciente da tecnologia para o gerenciamento. São Paulo: Editora Saraiva, 2004. 282 p.

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.

Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer Jr, E. F.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M.: The protein data bank: A computer-based archival file for macromolecular structures. *Journal of Molecular Biology* **1977**, 112, 535-542.

BioMat - Grupo de Física Teórica de Materiais e Biomoléculas. <http://www.df.ufscar.br/node/201>

Bioinformatics: [http://www.rcsb.org/pdb/static.do?p=general\\_information/about\\_pdb/](http://www.rcsb.org/pdb/static.do?p=general_information/about_pdb/)

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-[small pi] interactions in supramolecular hydrogels of aromatic carbohydrate amphiphiles. *Chemical Science* **2011**, 2, 1349-1355.

Burley, S. K.; Petsko, G. A.: Aromatic-aromatic interaction: a mechanism of protein structure stabilization. *Science* **1985**, 229, 23-28, 1985.

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)···π(aryl) synthon. *Coordination Chemistry Reviews* **2012**, 256, 412-438.

Cauët, E.; Rومان, M.; Wintjens, R.; Liévin, J.; Biot, C.. Histidine–Aromatic Interactions in Proteins and Protein–Ligand Complexes: Quantum Chemical Study of X-ray and Model Structures. *Journal of Chemical Theory and Computation*, v. 1, n. 3, p. 472-483, **2005**.

Contributors, U. EDSAC, Electronic Delay Storage Automatic Calculator. **2009**.  
<http://uncyclopedia.wikia.com/wiki/EDSAC>

Contributors, W. EDSAC, Electronic Delay Storage Automatic Calculator. **2012a**.  
[http://en.wikipedia.org/wiki/Electronic\\_Delay\\_Storage\\_Automatic\\_Calculator](http://en.wikipedia.org/wiki/Electronic_Delay_Storage_Automatic_Calculator)

Contributors, W. Ficheiro: X-ray diffraction.png – Wikipédia, a enciclopédia livre. **2012b**.  
[http://pt.wikipedia.org/wiki/Ficheiro:X\\_ray\\_diffraction.png](http://pt.wikipedia.org/wiki/Ficheiro:X_ray_diffraction.png)

Contributors. History of Jmol Development. **2012**.

<http://jmol.sourceforge.net/history/#Jmol> - an open source replacement for the Chime

Contributors. History of Jmol Development. **2012**.

<http://jmol.sourceforge.net/history/#Jmol> - an open source replacement for the Chime plugin

Cranswick , L. M. D.: Busting out of crystallography's Sisyphean prison: from pencil and paper to structure solving at the press of a button: past, present and future of crystallographic software development, maintenance and distribution. *Crystallography Journals Online Acta Crystallographica Section A- Foundations of Crystallography* **2008**, 65-87.

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.

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

Drenth, J.: Principles of Protein X-Ray Crystallography. Groningen: Springer Science+Business Media, 2007. 332 p.

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.; Meng, J.-Z.; Liao, S.-M.; Huang, R.-B.: Empirical formulation and parameterization of cation- $\pi$  interactions for protein modeling. *Journal of Computational Chemistry* **2012**, 33, 153-162.

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

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

Ferrin, T. E.; Huang, C. C.; Jarvis, L. E.; Langridge, R.: The MIDAS display system. *Journal of Molecular Graphics* **1988**, 6, 13-27.

Fritts, L. E.; Schwind, M. M.: Using the Wiswesser line notation (WLN) for online, interactive searching of chemical structures. *Journal of Chemical Information and Computer Sciences* **1982**, 22, 106-109.

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

Herráez, A.: Biomolecules in the computer: Jmol to the rescue. *Biochemistry and Molecular Biology Education* **2006**, 34, 255-261.

Howard, A. E.; Kollman, p. A. An analysis of current methodologies for conformational searching of complex molecules. *Journal of Medicinal Chemistry* **1988**, 31, 1669-1675.

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.

IBM: Specifications for The IBM Mathematical Formula Translater System. New York, 1954. 29 p.

Imai, Y. N.; Inoue, Y.; Nakanishi, I.; Kitaura, K.: Amide- $\pi$  interactions between formamide and benzene. *Journal of Computational Chemistry* **2009**, 30, 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.

JMOL. Jmol: an open-source Java viewer for chemical structures in 3D.  
<http://jmol.sourceforge.net/>

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.

Katz, L.; Levinthal, C.: Interactive Computer Graphics and Representation of Complex Biological Structures. *Annual Review of Biophysics and Bioengineering* **1972**, 1, 465-504.

Kendrew, J. C.; Bodo, G.; Dintzis, H. M.; Parrish, R. G.; Wyckoff, H.; Phillips, D. C.: A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis. *Nature* **1958**, 181, 662-666.

Khavasi, H. R.; Azizpoor Fard, M.:  $\pi$ - $\pi$  Interactions Affect Coordination Geometries. *Crystal Growth & Design* 2010, 10, 1892-1896.

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.

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.

LaCrEMM - Laboratório de Cristalografia Estereodinâmica e Modelagem Molecular.  
<http://www.quimica.ufscar.br/>

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

Leithold, L.: O cálculo com geometria analítica. [The Calculus with analytic geometry]. Cyro de Carvalho Patarra (Trad.). 3 ed. São Paulo: Harbra, c1994. v.1. 685 p.

Levinthal, C.: Molecular model-building by computer. *Scientific American* **1966**, 214, 42-52.

Lipschutz, S.; Lipson, M. L.: Teoria e Problemas de Matemática Discreta. 2 ed. Porto Alegre: Bookman, **2004**. 512 p.

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)](2) 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 Americal Chemical Society* **2004**, 126, 13659-13669.

Marsili, S.; Chelli, R.; Schettino, V.; Procacci, P.: Thermodynamics of stacking interactions in proteins. *Physical Chemistry Chemical Physics* **2008**, 10, 2673-2685.

Martz, E. (2004). Chime Version History.

<http://www.umass.edu/microbio/chime/chimvers.htm>

Martz, E. (2012a). Early Protein Crystallography.

[http://www.umass.edu/microbio/rasmol/1st\\_xtls.htm](http://www.umass.edu/microbio/rasmol/1st_xtls.htm)

Martz, E. (2012b). It's About Chime.

[http://www.umass.edu/microbio/chime/abtchime.htm#brief\\_history](http://www.umass.edu/microbio/chime/abtchime.htm#brief_history)

Martz, E. (1996). RasMol Quick Start.

<http://www.umass.edu/microbio/rasmol/rasquick.htm#mouse>

Martz, E.; Francoeur, E. (1997). History of Visualization of Biological Macromolecules. University of Massachusetts. <http://www.umass.edu/microbio/rasmol/history.htm>

RasMol: <http://www.umass.edu/microbio/rasmol/index2.htm>

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.

Meyer, E. F.: The first years of the Protein Data Bank. *Protein Science* **1997**, 6, 1591-1597.

Microsoft, Microsoft Corporation (**2012a**). Software, Smartphones, Online, Games, Cloud Computing, IT Business Technology, Downloads.

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

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

Minoux, H.; Chipot, C.: Cation- $\pi$  Interactions in Proteins: Can Simple Models Provide an Accurate Description? *Journal of the American Chemical Society* **1999**, *121*, 10366-10372.

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.

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.

Lehninger, A.L.; Nelson, D.L.; Cox, M.M.: Lehninger principios de bioquímica. [Lehninger principles of biochemistry]. Arnaldo Antonio Simoes (Coord.); Wilson Roberto Navega Lodi (Coord.). Arnaldo Antonio Simoes (Trad.). 3 ed. Sao Paulo: SARVIER, **2002**. 975 p.

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.

Protein Data Bank (PDB): [www.rcsb.org/pdb](http://www.rcsb.org/pdb)



PDBsum: <http://www.ebi.ac.uk/thornton-srv/databases/pdbsum/>

RasMol – University Of Massachusetts. RasMol Classic Site. 1995.  
<http://www.umass.edu/microbio/rasmol/index2.htm>

Ricarte, I. L. M.: Faculdade de Engenharia Elétrica e de Computação EA 960 - Organização de Computadores. Campinas: UNICAMP **1999**.

Sacco, A. C. S.; Souza, A. L.; Martins, H. B.; Guarnieri, G. F.: Elaboração de um Sistema de Informação para o Controle de Presença e Certificação do Simpósio de Tecnologia da Fatec-Tatuí. Parte I: Influência da Metodologia na Implantação. *Revista Sapere* **2012**, 4, 1-7.

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.

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

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.

Silberschatz, A.; Korth, H. F.; Sudarshan, S.: Sistema de banco de dados. [Database system concepts]. Daniel Vieira (Trad.). Rio de Janeiro: Elsevier, **2006**. 781 p.

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*, v. 112, n. 23, p. 5286-5294, **2008**.

Sörme, P.; Arnoux, P.; Kahl-Knutsson, B.; Leffler, H.; Rini, J. M.; Nilsson, U. J.: Structural and Thermodynamic Studies on Cation- $\pi$  Interactions in Lectin-Ligand Complexes: High-Affinity Galectin-3 Inhibitors through Fine-Tuning of an Arginine-Arene Interaction. *Journal of the American Chemical Society* **2005**, 127, 1737-1743.

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

Sun, S. F.: Physical chemistry for macromolecules. Second. Hoboken, NJ, USA: John Wiley & Sons, Inc., **2004**. 578 p.

Sussman, J.L.; Lin, D.; Jiang, J.; Manning, N. O.; Prilusky, J.; Ritter, O.; Abola, E. E.: Protein Data Bank (PDB): Database of Three-Dimensional Structural Information of Biological Macromolecules. *International Union of Crystallography* **2012**, 54, 1078-1084.

Sussman, J. L.; Abola, E. E.; Lin, D.; Jiang, J.; Manning, N. O.; Prilusky, J.: The protein data bank. *Genetica* **1999**, 106, 149-158.

Tenenbaum, A. M.; Langsam, Y.; Augenstein, M. J.: Estruturas de dados usando C. [Data structures using C]. São Paulo: Pearson Makron Books, **1995**. 884 p.

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.: Gold[small pi] aryl interactions as supramolecular synthons. *CrystEngComm* **2009**, 11, 1176-1186.

Tiekink, E. R. T.; Zukerman-Schpector, J.: Pb••• $\pi$  Aryl Interactions as Supramolecular Synthons. *Australian Journal of Chemistry* **2010**, 63, 535-543.

Tometsko, A. M.: Computer approaches to protein structure I. Analysis of atomic distances. *Computers and Biomedical Research* **1970**, 3, 229-237.

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 Attraction and Directionality of the  $\pi/\pi$  Interaction: Model Chemistry Calculations of Benzene Dimer Interaction. *Journal of the American Chemical Society*, v. 124, n. 1, p. 104-112, **2001**.

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.

Turban, E. R.; McLean, E. C.; Wetherbe, J.: Tecnologia da informação para gestão: transformando os negócios na economia digital. [Information technology for management: transforming business in the digital economy]. Renato Schinke (Trad.). 3 ed. Porto Alegre: Bookman, **2004**. 660 p.

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.

Voet, J.; Hammons, J. H. (2005). A short history of visualizing structures in the PDB. [http://www.rcsb.org/pdb/general\\_information/news\\_publications/newsletters/2005q1/education\\_corner.html](http://www.rcsb.org/pdb/general_information/news_publications/newsletters/2005q1/education_corner.html)

Westbrook, J.; Ito, N.; Nakamura, H.; Henrick, K.; Berman, H. M.: PDBML: the representation of archival macromolecular structure data in XML. *Bioinformatics* **2005**, 21, 988-992.

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.

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.

Wouters, J.. Cation- $\pi$  ( $\text{Na}^+$ -Trp) interactions in the crystal structure of tetragonal lysozyme. *Protein Science*, v. 7, n. 11, p. 2472-2475, **1998**.

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.

Zukerman-Schpector, J.; Otero-de-la-Roza, A.; Luana, V.; Tiekink, E.: Supramolecular architectures based on As(lone pair)[small pi](aryl) interactions. *Chemical Communications* **2011**, 47, 7608-7610.

---

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.

# Anexos

---



<b>Anexo 3.1</b>	<i>1GRA MOL2 - MAPEAMENTO DE INTERAÇÕES <math>\pi</math></i>
<b>Anexo 3.2</b>	<i>DNA-SL 1BNA - MOL2 - H PROTEINAS ATOMOS-CENTROIDES</i>
<b>Anexo 3.3</b>	<i>DNA-SL 1BNA - MOL2 - ATOMOS-CENTROIDES</i>
<b>Anexo 3.4</b>	<i>DNA-LS 1VZK - MOL2 - H PROTEINAS ATOMOS-CENTROIDES</i>
<b>Anexo 3.5</b>	<i>DNA-LS 1VZK - MOL2 - ATOMOS-CENTROIDES</i>
<b>Anexo 3.6</b>	<i>DNA-LI 1G3X - MOL2 - H PROTEINAS ATOMOS-CENTROIDES</i>
<b>Anexo 3.7</b>	<i>DNA-LI 1G3X - MOL2 - ATOMOS-CENTROIDES</i>
<b>Anexo 3.8</b>	<i>Tabulação e Estatística DNA</i>
<b>Anexo 3.9</b>	<i>Estruturas de DNA com interações <math>\pi</math></i>
<b>Anexo 3.10</b>	<i>Estruturas de DNA com interações <math>\pi</math> sem as do ligante</i>
<b>Anexo 3.11</b>	<i>Estruturas de DNA com interações <math>\pi</math> por tipo com ligantes</i>
<b>Anexo 3.12</b>	<i>Estruturas de DNA com interações <math>\pi</math> por tipo sem ligantes</i>
<b>Anexo 3.13</b>	<i>Deslocamento das coordenadas dos centroides - base DNA sem ligante</i>
<b>Anexo 3.14</b>	<i>Estruturas de aminoácidos e bases do DNA com a identificação dos átomos</i>