

**Alessandra Renata Lente da Silva**

***Estudo Estrutural de Quinases Dependentes de Ciclinas por  
Métodos de Modelagem Molecular Comparativa***

Dissertação apresentada para obtenção do título de Mestre em Biofísica Molecular, área de concentração Biofísica Molecular do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista "Júlio de Mesquita Filho" –UNESP.

**Orientadora: Profa. Dra. Fernanda Canduri**

São José do Rio Preto  
Estado de São Paulo – Brasil  
Dezembro – 2006

*“A mente que se abre a uma nova idéia jamais retorna ao seu tamanho original.”*

**(Albert Einstein)**

***Dedico este trabalho a todas as pessoas que lutam por aquilo em que acreditam, e, mesmo passando por dificuldades, não desistem de lutar por seus objetivos, ou seja, dedico àqueles que verdadeiramente possam ser chamados de vencedores!***

## AGRADECIMENTOS

Em primeiro lugar agradeço a DEUS pela forte presença em minha vida, e, pela grande força proporcionada nos bons e maus momentos; agradeço por ter auxiliado a me tornar à pessoa que sou hoje, como : filha, esposa, mãe, mulher, profissional e cristã...

Em segundo lugar agradeço ao Prof. Dr. Walter Filgueira de Azevedo Júnior pelo exemplo de profissional que é, e, no qual tenho como parâmetro; pelos seus conselhos que direcionaram na minha formação acadêmica, e principalmente pelo grande apoio e auxílio na realização de um dos grandes sonhos de minha vida, o Mestrado...

Em terceiro, agradeço a Profa. Dra. Fernanda Canduri, pelo profissionalismo que possui, e, sob sua orientação, sempre me ensinou como devo percorrer o caminho do conhecimento e de como ser competente em minha área; agradeço pela grande e sincera amizade que me proporciona, e por sua compreensão nos momentos difíceis que passei, principalmente quando tive depressão ... dos quais sua ajuda foi imprescindível...

Agradeço ao Prof. Dr. Jorge Chahine pelas informações sobre assuntos relacionados ao programa de Pós-graduação...

Agradeço ao Prof(s). Dr(s). Fernanda, João, Marcelo e Walter pelos conhecimentos transmitidos, nas disciplinas que ministraram no programa de Pós-graduação, das quais tive que cumprir, para obtenção de créditos...

Agradeço ao Prof. Dr. Ramón Beltran Abrego, pela amizade ,e pela oportunidade de efetuar o Estágio de Docência, junto a sua disciplina: Física Geral e Experimental I, com a turma de Licenciatura em Matemática do período noturno, no qual me acrescentou muito como profissional, por suas dicas e conselhos...

Agradeço ao Prof. Dr. Júlio César Borges, pela contribuição dada a este trabalho, e por permitir que eu descobrisse a excelente pessoa que é...

Agradeço aos Prof.(s).Dr(s). Marcelo Andres Fossey e Ramón Beltran Abrego, por aceitarem o convite para compor minha banca de qualificação, e, pela valiosa contribuição de conselhos e ensinamentos para minha tese final...

Agradeço aos prof(s).Dr(s). Ramón Beltran Abrego, e Andréia Machado Leopoldino pelo aceite ao convite para compor minha banca de defesa ...

Agradeço aos funcionários do Departamento de Física: Ilva, Barbosa e Paulinho, por todas às vezes que precisei de ajuda, sempre estiveram prontos a me auxiliar...

Agradeço aos funcionários da Secretaria de Pós-graduação, principalmente, a funcionária Silvia, pelas dúvidas esclarecidas, sempre com muita gentileza, todas às vezes que solicitei; e a funcionária Rosemar, que num momento de grande aflição, contribuiu com valiosos conselhos e estímulos de coragem, trazendo esperança para algo que acreditava estar perdido...

Agradeço aos meus irmãos “postiços”, Cristiane e Marcos Michel, pela amizade, companheirismo, fidelidade e paciência, em todas às vezes que precisei de um ombro amigo...

Agradeço as amigas Ana Helena e Karisa, pela ajuda obtida no ensino de como trabalhar com as ferramentas de bioinformática, tendo deixado de lado seus afazeres para contribuir com meu conhecimento científico, e, me incentivando a nunca desanimar da luta...

Agradeço as amigas Helen e Lisandra, pela amizade oferecida, pois num momento de total solidão, estenderam suas mãos para mim, ensinando a “nunca julgar um livro por sua capa”, e assim fazendo, provaram que uma amizade sincera nunca acaba...

Agradeço aos colegas do Laboratório de Sistemas Biomoleculares, principalmente ao amigo Márcio, pelas conversas, risos e trocas de experiências que me proporcionou...

E por falar em amizade, agradeço a minha grande e inesquecível amiga, Adriana Permigliani, que me ensinou grandes lições de vida, e esteve presente em momentos difíceis...

Agradeço aos Prof(s). Drs(s). Fernando Cachucho da Silva e Plínio Delatorre, por serem exemplos de mestres e por terem me ensinado que a Ciência é algo espetacular...

Agradeço a todos os colegas e professores que passaram por minha vida escolar e que de alguma maneira contribuíram para minha formação...

Agradeço a meus pais Rubens e Sônia, que sempre foram exemplos de vida para minha formação moral e principalmente por me ensinarem que a maior riqueza que um pai pode deixar a seu filho, é o estudo, pois com ele, adquire-se o conhecimento, e este, ninguém rouba e nunca acaba, somente é acrescido...

Agradeço a meu irmão, que verdadeiramente posso chamar de “garoto prodígio”, pois aos vinte e seis anos já alcançou seu grande sonho, ser professor universitário de instituição federal; pelo exemplo de vida e companheirismo, só tenho a agradecer...

Agradeço as duas maiores riquezas de minha vida, as quais amo do fundo do meu coração e que são meus “pilares”, meu “chão”, as duas pessoas que me inspiram a sempre lutar, sem cessar: meu amor, meu amigo, meu esposo Ronaldo e minha grande paixão, meu fruto mais precioso, meu filho Miguel David...

Agradeço o apoio financeiro da CAPES pela bolsa de Mestrado, e a FAPESP pelo suporte técnico do Laboratório de Sistemas Biomoleculares...

Finalmente, agradeço a todos aqueles que estão agora lendo este trabalho...

# SUMÁRIO

|                                                                                    | Página |
|------------------------------------------------------------------------------------|--------|
| LISTA DE FIGURAS .....                                                             | ix     |
| LISTA DE TABELAS .....                                                             | xi     |
| LISTA DE ABREVIACÕES .....                                                         | xii    |
| RESUMO .....                                                                       | xiii   |
| ABSTRACT .....                                                                     | xiv    |
| 1. INTRODUÇÃO .....                                                                | 1      |
| 1.1. Ciclo Celular .....                                                           | 1      |
| 1.2. Sistema Controle do Ciclo Celular .....                                       | 2      |
| 1.3. Proteíno-quinases ativadas ciclicamente .....                                 | 3      |
| 1.4. Quinases Dependentes de Ciclinas .....                                        | 4      |
| 1.5. Funções das CDKs .....                                                        | 5      |
| 1.6. Sobre os Inibidores de CDKs .....                                             | 6      |
| 1.7. Propriedades dos Inibidores .....                                             | 7      |
| 1.8. Características dos inibidores extrínsecos .....                              | 8      |
| 1.9. CDKs 8 e 10 como alvos para desenho de inibidores .....                       | 10     |
| 1.10. Um pouco sobre a CDK2 (molde) utilizado nas modelagens das CDKs 8 e 10 ..... | 11     |
| 2. OBJETIVOS .....                                                                 | 13     |
| 3. METODOLOGIA E FERRAMENTAS .....                                                 | 14     |
| 3.1. Bioinformática .....                                                          | 14     |
| 3.2. Modelagem Molecular por Homologia ou Comparativa .....                        | 15     |
| 3.2.1. Procura e Seleção do Molde .....                                            | 18     |
| 3.2.2. Alinhamento e Construção do Modelo .....                                    | 18     |
| 3.2.3. Avaliação do Modelo .....                                                   | 19     |
| 3.3. Métodos de Avaliação .....                                                    | 20     |
| 4. ANÁLISE E RESULTADOS .....                                                      | 27     |

|                                                |    |
|------------------------------------------------|----|
| 4.1. Comparação das Seqüências Primárias ..... | 27 |
| 4.2. Qualidade do Modelo .....                 | 28 |
| 4.3. Descrição dos Modelos .....               | 34 |
| 4.4. Interação CDK: Inibidores .....           | 41 |
| 5. CONCLUSÃO .....                             | 52 |
| 6. REFERÊNCIAS BIBLIOGRÁFICAS .....            | 53 |

## LISTA DE FIGURAS

|                                                                                                                                                                                                                                                                                       | Página |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|
| <b>Figura 1-1</b> As quatro fases sucessivas de um ciclo celular .....                                                                                                                                                                                                                | 02     |
| <b>Figura 1-2</b> Dois pontos de checagem do ciclo celular .....                                                                                                                                                                                                                      | 03     |
| <b>Figura 1-3</b> As reações de fosforilação/defosforilação. Dois tipos de enzimas com funções opostas. Quinases e fosfatases regulam o nível de fosforilação de proteínas.....                                                                                                       | 04     |
| <b>Figura 1-4</b> Algumas CDKs e as respectivas ciclinas, relacionando-as com suas funções celulares ...                                                                                                                                                                              | 06     |
| <b>Figura 1-5</b> Diagrama de como é realizada a inibição do processo de ligação dos inibidores nos compostos de CDKs .....                                                                                                                                                           | 07     |
| <b>Figura 1-6</b> Flavopiridol .....                                                                                                                                                                                                                                                  | 09     |
| <b>Figura 1-7</b> Olomoucina .....                                                                                                                                                                                                                                                    | 09     |
| <b>Figura 1-8</b> Roscovitina .....                                                                                                                                                                                                                                                   | 10     |
| <b>Figura 1.9</b> A CDK 8 e a ciclina C são componentes do complexo da RNA polimerase II, funcionam como uma proteína quinase que fosforila o domínio C-terminal da subunidade maior da RNA polimerase II. A CDK10 está envolvida no ciclo celular, na transcrição da fase G2/M ..... | 11     |
| <b>Figura 1.10</b> Estrutura da CDK2 (PDB : 1HCL) .....                                                                                                                                                                                                                               | 12     |
| <b>Figura 1.11</b> Diagrama Esquemático do Garfo Molecular (DE AZEVEDO, 2002) .....                                                                                                                                                                                                   | 12     |
| <b>Figura 3-1</b> Interação das áreas que contribuem para a formação e desenvolvimento da bioinformática .....                                                                                                                                                                        | 14     |
| <b>Figura 3-2</b> Passos da modelagem molecular comparativa .....                                                                                                                                                                                                                     | 17     |
| <b>Figura 3-3</b> Parte de uma cadeia polipeptídica indicando os graus de liberdade torcional de cada unidade peptídica e os ângulos de rotação $\phi$ e $\varphi$ .....                                                                                                              | 21     |
| <b>Figura 3-4</b> Diagrama de Ramachandran .....                                                                                                                                                                                                                                      | 22     |
| <b>Figura 4-1</b> Figuras dos alinhamentos das seqüências primárias em (a) CDK8 e CDK2; em (b) CDK10 e CDK2 .....                                                                                                                                                                     | 28     |
| <b>Figura 4-2</b> Diagramas de Ramachandran das (a) CDK8:ATP, (b) CDK8:flp, (c) CDK8:olo, (d) CDK8:rsc .....                                                                                                                                                                          | 29     |
| <b>Figura 4-3</b> Diagramas de Ramachandran das (a) CDK10:ATP, (b) CDK10:flp, (c) CDK10:olo, (d) CDK10:rsc .....                                                                                                                                                                      | 30     |
| <b>Figura 4-4</b> Estrutura da CDK8 complexada com ATP. No detalhe, a interação do ATP com o sítio ativo da proteína.....                                                                                                                                                             | 34     |
| <b>Figura 4-5</b> Estrutura da CDK8 complexada com flavopiridol. No detalhe, a interação do flavopiridol com o sítio ativo da proteína.....                                                                                                                                           | 35     |
| <b>Figura 4-6</b> Estrutura da CDK8 complexada com olomoucina. No detalhe, a interação da olomoucina com o sítio ativo da proteína.....                                                                                                                                               | 35     |

|                                                                                                                                                                                             |    |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| <b>Figura 4-7</b> Estrutura da CDK8 complexada com roscovitina. No detalhe, a interação da roscovitina com o sítio ativo da proteína .....                                                  | 36 |
| <b>Figura 4-8</b> Estrutura da CDK10 complexada com ATP. No detalhe, a interação do ATP com o sítio ativo da proteína .....                                                                 | 36 |
| <b>Figura 4-9</b> Estrutura da CDK10 complexada com flavopiridol. No detalhe, a interação do flavopiridol com o sítio ativo da proteína .....                                               | 37 |
| <b>Figura 4-10</b> Estrutura da CDK10 complexada com olomoucina. No detalhe, a interação da olomoucina com o sítio ativo da proteína .....                                                  | 37 |
| <b>Figura 4-11</b> Estrutura da CDK10 complexada com roscovitina. No detalhe, a interação da roscovitina com o sítio ativo da proteína .....                                                | 38 |
| <b>Figura 4-12</b> Sobreposição da CDK8 com seu molde CDK2, em (a) CDK8:ATP com CDK2:ATP; em (b) CDK8:flp com CDK2:flp; (c) CDK8:olo com CDK2:olo e em (d) CDK8:rsc com CDK2:rsc .....      | 39 |
| <b>Figura 4-13</b> Sobreposição da CDK10 com seu molde CDK2, em (a) CDK10:ATP com CDK2:ATP; em (b) CDK10:flp com CDK2:flp; (c) CDK10:olo com CDK2:olo e em (d) CDK10:rsc com CDK2:rsc ..... | 41 |
| <b>Figura 4-14</b> Sobreposição dos sítios de ligação da CDK2 em preto e da CDK8 em vermelho, na presença de ATP .....                                                                      | 42 |
| <b>Figura 4-15</b> Sobreposição dos sítios de ligação da CDK2 em preto e da CDK8 em vermelho, na presença de flavopiridol .....                                                             | 43 |
| <b>Figura 4-16</b> Sobreposição dos sítios de ligação da CDK2 em preto e da CDK8 em vermelho, na presença de Olomoucina.....                                                                | 44 |
| <b>Figura 4-17</b> Sobreposição dos sítios de ligação da CDK2 em preto e da CDK8 em vermelho, na presença de Roscovitina.....                                                               | 45 |
| <b>Figura 4-18</b> Sobreposição dos sítios de ligação da CDK2 em preto e da CDK10 em azul, na presença de ATP .....                                                                         | 46 |
| <b>Figura 4-19</b> Sobreposição dos sítios de ligação da CDK2 em preto e da CDK10 em azul, na presença de flavopiridol.....                                                                 | 47 |
| <b>Figura 4-20</b> Sobreposição dos sítios de ligação da CDK2 em preto e da CDK10 em azul, na presença de Olomoucina.....                                                                   | 48 |
| <b>Figura 4-21</b> Sobreposição dos sítios de ligação da CDK2 em preto e da CDK10 em azul, ligados na presença de Roscovitina.....                                                          | 49 |
| <b>Figura 4-22</b> Potenciais Eletrostáticos dos quatro complexos da CDK8.....                                                                                                              | 50 |
| <b>Figura 4-23</b> Potenciais Eletrostáticos dos quatro complexos da CDK10.....                                                                                                             | 51 |

## LISTA DE TABELAS

|                                                                                                                                                                                          | Página |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|
| <b>Tabela 1-1</b> IC <sub>50</sub> para diversos inibidores de CDKs .....                                                                                                                | 08     |
| <b>Tabela 4-1</b> Resultados das análises do Diagrama de Ramachandran dos modelos (CDK8 complexada com: ATP, flp, olo, rsc) e de seu moldes (CDK2 complexada com: ATP, flp, olo, rsc) .  | 29     |
| <b>Tabela 4-2</b> Resultados das análises do Diagrama de Ramachandran dos modelos (CDK10 complexada com: ATP, flp, olo, rsc) e de seu moldes (CDK2 complexada com: ATP, flp, olo, rsc) . | 30     |
| <b>Tabela 4-3</b> Resultado dos valores do G-Factor dos modelos de CDK8 com ATP, flp, olo e rsc. ....                                                                                    | 31     |
| <b>Tabela 4-4</b> Resultado dos valores do G-Factor dos modelos de CDK 10 com ATP, flp, olo e rsc ...                                                                                    | 32     |
| <b>Tabela 4-5</b> Resultado dos valores do VERIFY 3D dos modelos de CDK8 com seus respectivos complexos, os moldes de CDK2 estão entre parênteses .....                                  | 32     |
| <b>Tabela 4-6</b> Resultado dos valores do VERIFY 3D dos modelos de CDK 10 com seus respectivos complexos, os moldes de CDK2 estão entre parênteses .....                                | 33     |
| <b>Tabela 4-7</b> Resultado dos valores do RMSD da geometria ideal dos modelos de CDK8 .....                                                                                             | 33     |
| <b>Tabela 4-8</b> Resultado dos valores do RMSD da geometria ideal dos modelos de CDK10 .....                                                                                            | 34     |
| <b>Tabela 4-9</b> Valores de RMSD da sobreposição para a CDK8 .....                                                                                                                      | 38     |
| <b>Tabela 4-10</b> Valores de RMSD da sobreposição para a CDK10 .....                                                                                                                    | 40     |
| <b>Tabela 4-11</b> Ligações de Hidrogênio entre os resíduos de aa da CDK2 e CDK8 e os átomos do ATP                                                                                      | 42     |
| <b>Tabela 4-12</b> Ligações de Hidrogênio entre os resíduos de aa da CDK2 e CDK8 e os átomos do Flavopiridol .....                                                                       | 43     |
| <b>Tabela 4-13</b> Ligações de Hidrogênio entre os resíduos de aa da CDK2 e CDK8 e os átomos da Olomoucina .....                                                                         | 44     |
| <b>Tabela 4-14</b> Ligações de Hidrogênio entre os resíduos de aa da CDK2 e CDK8 e os átomos da Roscovitina .....                                                                        | 45     |
| <b>Tabela 4-15</b> Ligações de Hidrogênio entre os resíduos de aa da CDK2 e CDK10 e os átomos do ATP .....                                                                               | 46     |
| <b>Tabela 4-16</b> Ligações de Hidrogênio entre os resíduos de aa da CDK2 e CDK10 e os átomos do Flavopiridol .....                                                                      | 47     |
| <b>Tabela 4-17</b> Ligações de Hidrogênio entre os resíduos de aa da CDK2 e CDK10 e os átomos da Olomoucina .....                                                                        | 48     |
| <b>Tabela 4-18</b> Ligações de Hidrogênio entre os resíduos de aa da CDK2 e CDK10 e os átomos da Roscovitina .....                                                                       | 49     |
| <b>Tabela 4-19</b> Tabela com valores da área de contato para cada tipo de inibidor e sua respectiva proteína .....                                                                      | 49     |

## LISTA DE ABREVIACÕES

|                             |                                                                    |
|-----------------------------|--------------------------------------------------------------------|
| <b>CDKs</b>                 | Cyclin-dependent-kinases ( Quinases dependentes de ciclinas )-     |
| <b>ATP</b>                  | Adenosina trifosfato                                               |
| <b>M</b>                    | Mitose                                                             |
| <b>G1</b>                   | Fase que compreende o fim da mitose e o início da síntese          |
| <b>S</b>                    | Síntese                                                            |
| <b>G2</b>                   | Fase que compreende o término da síntese e o início da mitose      |
| <b>DNA</b>                  | Ácido desoxirribonucleico                                          |
| <b>TFIIH</b>                | Fator de iniciação da transcrição IIH                              |
| <b>RNA</b>                  | Ácido ribonucleico                                                 |
| <b>P-TEFb</b>               | Fator de transcrição que estimula a elongação da RNA Polimerase II |
| <b>IC<sub>50</sub></b>      | Concentração do poder de inibição ou atividade por 50%             |
| <b>kDa</b>                  | kilo Dalton                                                        |
| <b>PDB</b>                  | Protein Data Bank ( Banco de Dados de Proteínas )                  |
| <b>RMN</b>                  | Ressonância Magnética Nuclear                                      |
| <b>3D</b>                   | Tridimensional                                                     |
| <b>RMSD</b>                 | Root Mean Squared Deviation ( Desvio Médio Quadrático )            |
| <b>C<math>\alpha</math></b> | Carbono alfa                                                       |
| <b>CCP4</b>                 | Collaborative Computational Project nº 4 ( Projeto nº 4 )          |
| <b><math>\Phi</math></b>    | PHI                                                                |
| <b><math>\phi</math></b>    | phi                                                                |
| <b><math>\psi</math></b>    | psi                                                                |
| <b>Å</b>                    | Angstrom                                                           |
| <b>C<math>\beta</math></b>  | Carbono beta                                                       |
| <b>pKa</b>                  | -logKa (constante de dissociação)                                  |
| <b>gaps</b>                 | lacunas                                                            |
| <b>flp</b>                  | flavopiridol                                                       |
| <b>olo</b>                  | olomoucina                                                         |
| <b>rsc</b>                  | roscovitina                                                        |

## RESUMO

As CDKs têm sido extensivamente estudadas, porque exercem papéis essenciais na regulação da proliferação celular, nos processos neuronais e transcricionais. A progressão do ciclo celular é firmemente controlada pela atividade das CDKs. Este projeto de Mestrado tem como objetivo estudar a interação das proteínas CDK8 e CDK10 com vários inibidores, e estabelecer as bases estruturais para inibição das mesmas. Foram realizadas modelagens moleculares dos complexos binários CDKs:Inibidores, utilizando como molde a estrutura da CDK2 complexada com os mesmos ligantes: ATP, flavopiridol, olomoucina e roscovitina. Para realização dessas modelagens foi utilizado o programa Parmodel, que é uma implementação paralela do programa MODELLER. Espera-se que os modelos computacionais produzam avanços no entendimento de suas estruturas, destacando-se seus sítios de ligações, para um possível estudo sobre inibidores específicos para estas CDKs.

## **ABSTRACT**

The CDKs have been extensively studied because of their essential role in the regulation of cell proliferation, in the neuronal process and transcription. Cell cycle progression is tightly controlled by the activity of CDKs. This master project aimed to study the interaction of the CDK8 and CDK10 with several inhibitors in order to establish the structural bases for its inhibition. Were carried out the molecular modeling of the binary complexes CDKs:Inhibitors, using as template the CDK2 structure with the following ligands and inhibitors: ATP, Flavopiridol, Olomoucine and Roscovitine. For the achievement of those modeling was utilized the Parmodel program which is a parallel implementation of the MODELLER program. One expects that the computational models produce advances in the agreement structures, showing their binding sites, for a possible study about specific inhibitors against CDKs.

# 1. INTRODUÇÃO

## 1.1. Ciclo Celular

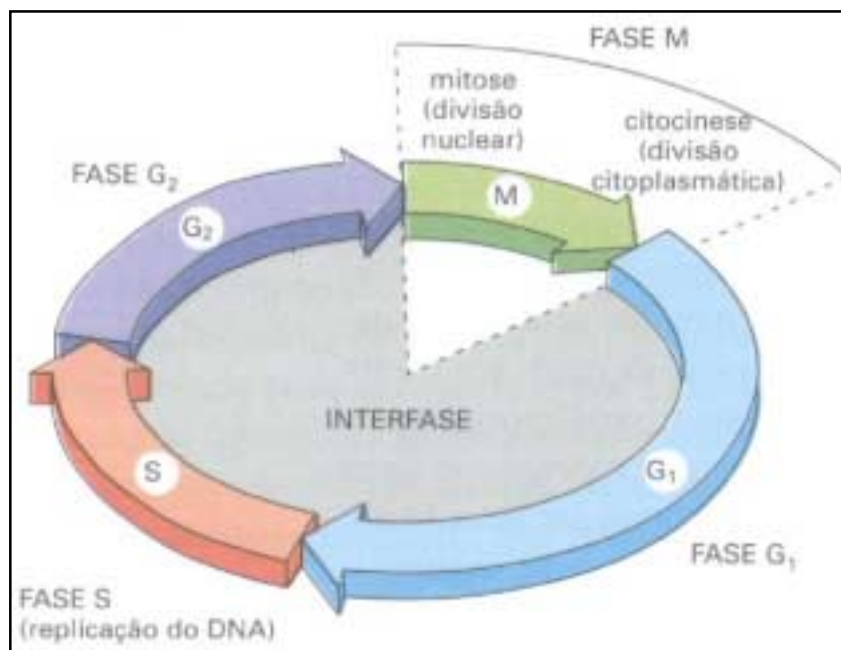
Uma célula se reproduz por uma seqüência ordenada de eventos que duplicam seus componentes e depois a divide em duas. Este ciclo de duplicação e divisão, é conhecido como **ciclo celular** (ALBERTS et al., 1999). O ciclo celular constitui o processo básico da gênese de novas células (JUNQUEIRA & CARNEIRO, 1997). Ele representa o mecanismo essencial pelo qual todos os seres vivos se reproduzem (ALBERTS et al., 1999).

Todos os organismos vivos, das bactérias unicelulares aos mamíferos multicelulares, são produtos de repetidas rodadas de crescimento celular e divisão desde o início da vida há mais de três bilhões de anos (ALBERTS et al., 1999).

Este ciclo de duplicação e divisão celular é dividido em quatro estágios ou fases (ALBERTS et al., 1999) (Figura 1-1).

Os dois eventos mais dramáticos acontecem quando o núcleo se divide, processo denominado mitose (M), e quando a célula divide-se em duas, processo denominado citocinese. Estes dois processos compõem a fase M do ciclo celular (ALBERTS et al., 1999).

O período entre uma fase M e a próxima é denominado interfase, que divide as três fases remanescentes do ciclo celular: fase  $G_1$ , fase S (síntese), e fase  $G_2$  (Figura 1-1). Durante a fase  $G_1$  inicia-se o intervalo entre o término da fase M e o início de S; na fase S, ocorre a replicação do DNA nuclear; e a fase  $G_2$  compreende o término da fase S e o início de M. Lembrando-se que as fases  $G_1$  e  $G_2$  compreendem “um período adicional” que as células possuem para crescer e duplicar suas organelas citoplasmáticas (ALBERTS et al., 1999).



**Figura 1-1** As quatro fases sucessivas de um ciclo celular (ALBERTS, et al., 1999).

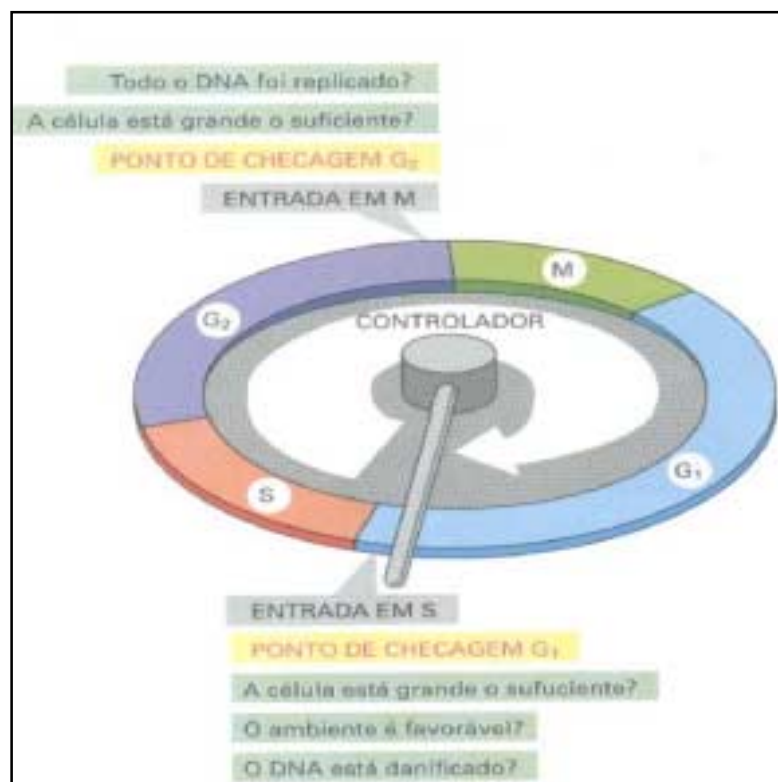
## 1.2. Sistema Controle do Ciclo Celular

Dois tipos de mecanismos são necessários no ciclo celular: um para produzir os novos componentes da célula em crescimento e outro para organizar os componentes nos seus devidos locais e dividi-los apropriadamente quando a célula divide-se em duas (ALBERTS et al., 1999).

O Sistema de Controle do Ciclo Celular assegura a progressão correta pelo ciclo celular, por meio da regulação da maquinaria citada acima (ALBERTS et al., 1999).

Esse Sistema de Controle funciona em uma sequência ordenada de eventos, e controla todo o processo de etapas do ciclo celular, através de seus “freios moleculares” que param em vários pontos para checagem (Figura 1-2). Devido a esse trabalho, ele não passa para a próxima etapa do ciclo celular, sem haver terminado a etapa anterior.

Os pontos de checagem são importantes porque são locais no ciclo celular, onde o Sistema Controle pode ser regulado por meio de sinais oriundos de outras células (fatores de crescimento, ou outras moléculas sinalizadoras extracelulares), que podem inibir ou promover a proliferação celular (ALBERTS et al., 1999).



**Figura 1-2** Dois pontos de checagem do ciclo celular (ALBERTS, et al., 1999).

### 1.3. Proteína-quinases ativadas ciclicamente

A fosforilação e desfosforilação são meios que as células utilizam para alterar atividades de proteínas. Esses meios são executados por um grupo específico de proteínas, as proteína-quinases (proteínas-chaves).

O papel das proteína-quinases, está em catalisar a transferência de um grupo fosfato do ATP a uma cadeia lateral de um aminoácido particular presente na proteína-alvo. Pode-se reverter o processo de fosforilação, pela remoção do grupo fosfato, desfosforilação, pelas proteína-fosfatases (Figura 1-3).

As proteína-quinases formam o Sistema Controle, e, estão presentes durante todo o ciclo celular (ALBERTS et al., 1999). Elas são ativadas em determinados momentos do ciclo, e após efetuar seu papel, são rapidamente desativadas. O grupo protéico responsável pela ativação ou desativação das proteína-quinases são as ciclinas.

As ciclinas recebem esse nome, devido sua concentração variar de maneira cíclica durante o ciclo celular. Elas possuem atividade enzimática, quando ligadas as quinases do ciclo celular, tornando essas também enzimaticamente ativas.

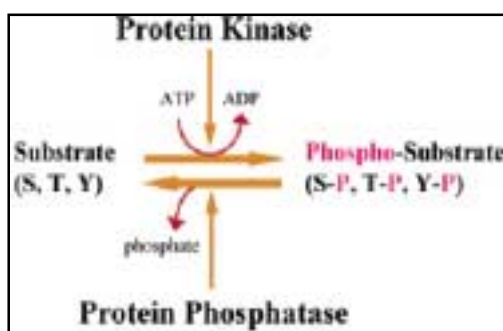
As proteínas quinases que só apresentam atividade na presença de ciclinas são denominadas quinases dependentes de ciclinas (“Cyclin-dependent Kinases” - CDKs).

#### 1.4. Quinases Dependentes de Ciclinas

O Sistema Controle possui função central na regulação do número celular nos tecidos dos organismos, na replicação do DNA, na mitose e citocinese (ALBERTS et al., 1999). E como visto, o Sistema Controle compreende duas famílias de proteínas, que representam chaves para seu funcionamento (BARVIAN et al., 2000). Uma dessas famílias são representadas pelas Quinases Dependentes de Ciclina (DE AZEVEDO et al., 1996). As CDKs são subunidades catalíticas de complexos heterodiméricos (GOULD, 2000), que recebem esse nome devido ao fato de não possuir atividade, a menos que estejam associadas com um membro da família das ciclinas (DE AZEVEDO et al., 1996, GOULD, 2000) denominadas proteínas ativadoras (DE AZEVEDO et al., 1996).

As CDKs são ativadas momentaneamente em estágios específicos do ciclo celular, e sua ativação ou inativação dispara os próximos eventos do ciclo celular (GOULD, 2000). Cada subunidade catalítica da CDK pode associar-se a diferentes ciclinas, e estas determinam quais serão fosforiladas pelo complexo CDK-ciclina (LODISH et al., 1995).

As CDKs respondem pela fosforilação dos resíduos de Serina, Treonina e Tirosina.



**Figura 1-3** As reações de fosforilação/defosforilação. Dois tipos de enzimas com funções opostas. Quinases e fosfatases regulam o nível de fosforilação de proteínas (MEIJER et al., 2002).

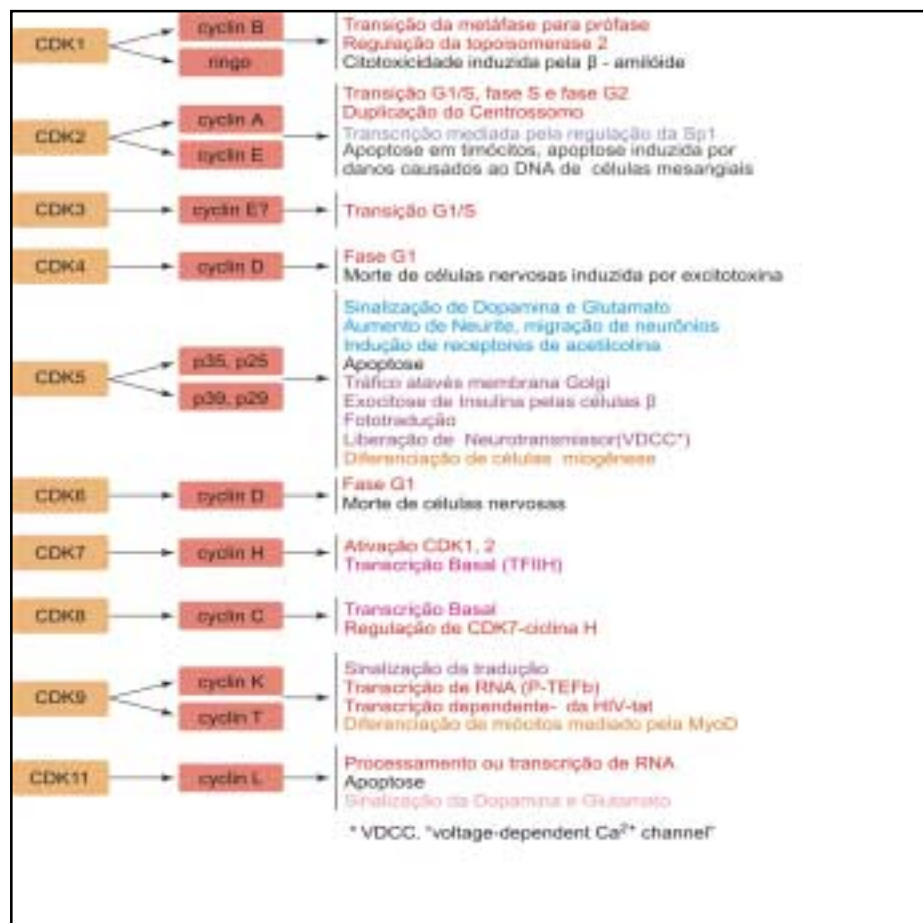
Este é o principal mecanismo intracelular pela qual as proteínas estruturais e as enzimas são reguladas a nível pós-transcricional. A fosforilação de uma proteína resulta no balanço entre as atividades das quinases e fosfatases (MEIJER et al., 2002) (Figura 1-3).

### 1.5. Funções das CDKs (KNOCKAERT et al., 2002)

Existem ao todo 13 CDKs codificadas no genoma humano, e essas, desempenham funções essenciais, no controle do ciclo celular, nos processos neuronais e transcripcionais, bem como na diferenciação e morte celular:

- Ciclo de Divisão Celular : a CDK1 é requerida em G2 e na transição G2/M, a CDK2 é importante em G1 e no início da fase S, e a CDK4 no início de G1;
- Transcrição : são quatro CDKs que regulam o processo de transcrição. A CDK7-ciclina H, é um componente do fator de transcrição TFIIH, a mesma CDK7-ciclina H e a CDK8-ciclina C, fosforilam uma ampla subunidade da RNA polimerase II requerida para a elongação da transcrição, a CDK9-ciclina T é um componente do fator de transcrição P-TEFb, e a CDK11 é importante na regulação do processamento e transcrição do RNA, quando associado à ciclina L;
- Funções Neuronais: a CDK5 desempenha inúmeras funções no sistema nervoso, incluindo desenvolvimento de neurites, migração neuronal, via do receptor metabotrópico de glutamato e via de sinalização da dopamina, e a CDK11, nas vias de sinalização da dopamina e do glutamato;
- Diferenciação: ativação ou inibição de CDKs exercem papéis centrais na diferenciação. Efetores positivos das CDK5 e CDK9, e efetores negativos da CDK2 na diferenciação de células musculares, e a re-diferenciação de células leucêmicas por inibição seletiva da CDK2 e CDK6;
- Morte Celular: a CDK1, CDK4 e CDK6, exercem função essencial na morte de células neuronais e as CDKs 5 e 11 também participam da apoptose;
- Outras funções: CDKs estão envolvidas no tráfego através da membrana do Complexo de Golgi, na exocitose de insulina pela células  $\beta$  do pâncreas e na regulação da enzima fosfodiesterase retinal.

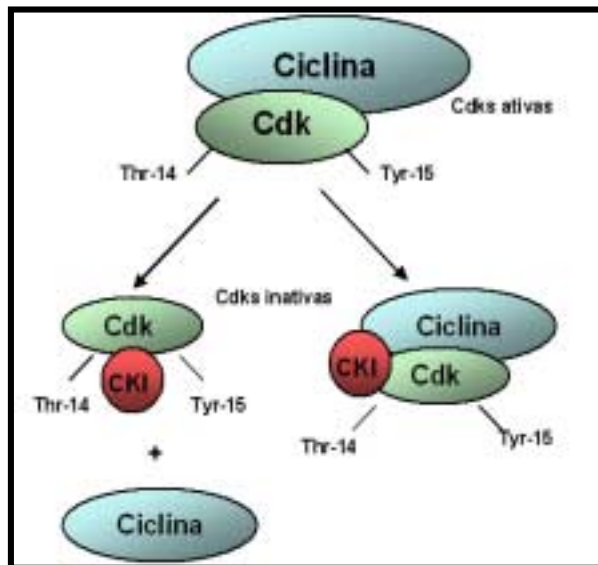
A Figura 1-4 apresenta um resumo das funções desempenhadas pelas CDKs:



**Figura 1-4** Algumas CDKs e as respectivas ciclinas, relacionando-as com suas funções celulares. (KNOCKAERT et al., 2002).

## 1.6. Sobre os Inibidores de CDKs

Os inibidores intrínsecos de CDKs, as CKIs, são proteínas que bloqueiam a atividade dos complexos de CDK-ciclina (ALBERTS *et al.*, 1999). Elas ligam-se às CDKs, ou aos dímeros CDKs-ciclina, e nas duas condições as quinases são inativadas (Figura 1-5) (GOULD, 2000).



**Figura 1-5** Diagrama de como é realizada a inibição do Processo de ligação dos inibidores nos complexos CDKs (GOULD,2000).

Quando se trata de inibidores extrínsecos de CDKs, fica demonstrado, que a seletividade é uma questão chave. Nenhum inibidor, até hoje, foi descoberto, evidenciando ser mais seletivo para uma CDK do que para outra (KNOCKAERT, et al., 2002). Os modelos estruturais de CDKs podem providenciar informações sobre desenhos de inibidores seletivos para CDKs. Mais de 50 inibidores farmacológicos de CDKs foram descritos, alguns com atividade potencial antitumoral (câncer) (DAÍ & GRANT, 2003).

### 1.7. Propriedades dos inibidores extrínsecos

São três as propriedades que tornam os inibidores extrínsecos de CDKs, potenciais agentes atrativos anti tumores:

- 1- são potentes agentes anti-proliferativos, interrompendo células em G1 (SONI et al., 2001, KNOCKAERT et al., 2002) ou G2/M (KNOCKAERT et al., 2002);
- 2- disparam apoptose, sozinhas ou em combinação com outros tratamentos (EDAMATSU et al., 2000);
- 3- em alguns exemplos, contribuem para a diferenciação celular (KNOCKAERT et al., 2002);

## 1.8. Características dos inibidores extrínsecos

Os inibidores de CDKs possuem baixa massa molecular (<600 da), são uniformes, heterocíclicos, hidrofóbicos, competitivos, e análogos de ATP (pois competem com ele para ligar-se ao sítio ativo) e ligam-se (muito mais) por interação hidrofóbica e pontes de hidrogênio às quinases (KNOCKAERT et al., 2002).

Neste trabalho, serão abordados os seguintes inibidores de CDKs: Flavopiridol, Olomoucina e Roscovitina, baseados nos valores de IC<sub>50</sub> (Concentração do poder de inibição ou atividade por 50%) para a CDK2 (que será utilizada como molde) (Tabela 1-1), e baseado também na identidade de seqüências entre as proteínas alvo e o molde.

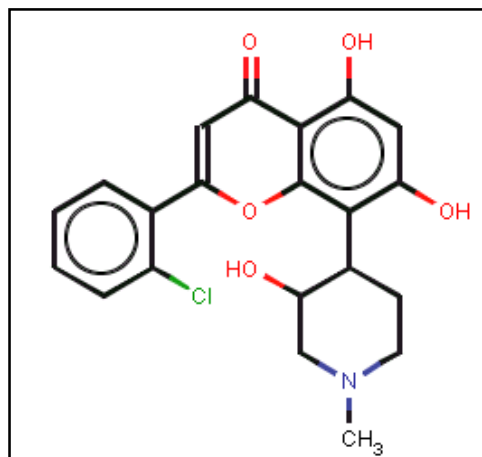
**Tabela 1-1** IC<sub>50</sub> para diversos inibidores de CDKs (todos os valores em  $\mu\text{M}$ ).

| Proteína       | CDK1 | CDK2 | CDK4  | CDK5 | CDK6 | CDK7 |
|----------------|------|------|-------|------|------|------|
| <b>Ligante</b> |      |      |       |      |      |      |
| Flavopiridol   | 0,4  | 0,1  | 0,4   | ND   | ND   | 0,3  |
| Olomoucina     | 7    | 7    | >1000 | 3    | >250 | ND   |
| Roscovitina    | 0,65 | 0,7  | >100  | 0,16 | >100 | ND   |

- ND: Não Determinado

### Flavopiridol

Flavopiridol, [ cis-5,7-diidroxi-2-( 2-chorofenil )-8-[ 4-3-hidroxi-1-metil )-piperidinil]- 1-benzopirano-4-1 ] é um bom inibidor de CDK, devido a suas propriedades antiproliferativas. É uma molécula pequena semi-sintética, que foi identificado na casca de árvore da planta *Dysoxylum binectariferum* (MANHANI et al,2005). Ele é um potencial agente terapêutico anticâncer, composto já em testes clínicos de fase I e II. O tratamento com Flavopiridol resulta no bloqueio da progressão do ciclo celular, promovendo diferenciação, e induzindo à apoptose em vários tipos de células tumorais (SENDEROWICZ & SAUSVILLE, 2000, DE AZEVEDO et al., 2002).

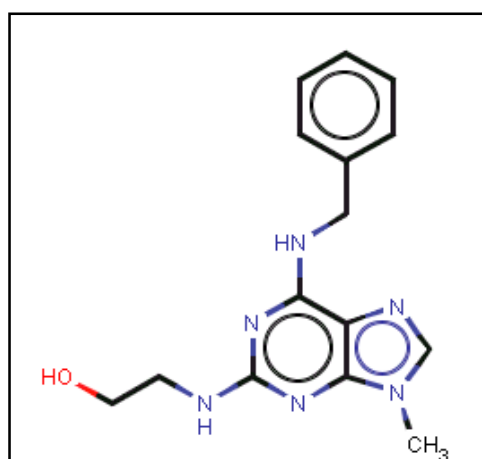


**Figura 1-6** Flavopiridol, (MEIJER et al., 2002; SIELECKI et al., 2000).

## Olomoucina

A Olomoucina (2-(2-hidroxietilamino)-6-benzilamino-9-metilpurina), assim como os demais, é um inibidor competitivo. O anel de purina da Olomoucina ocupa aproximadamente a mesma posição do anel do ATP (FURET et al., 2000).

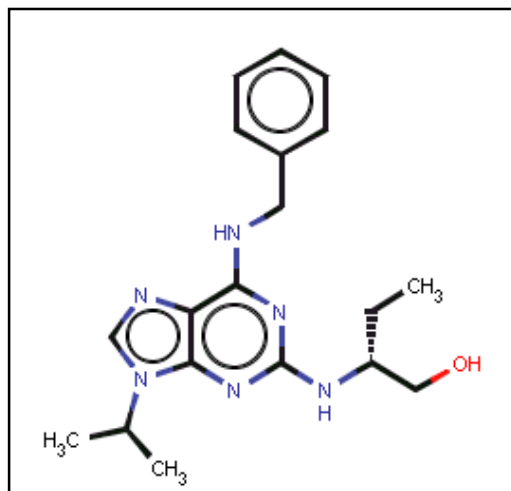
Estudos, mostraram que altas doses de olomoucina são capazes de bloquear completamente a progressão do ciclo celular e induzir a apoptose (SCHUTTE et al., 1997).



**Figura 1-7** Olomoucina, (MEIJER et al., 2002; SIELECKI et al., 2000).

## Roscovitina

A Roscovitina é um inibidor que potencialmente inibe CDKs, com valores de  $IC_{50}$  em torno de 0,16  $\mu$ M (para CDK5) e 100  $\mu$ M (para CDK4 e CDK6) (MEIJER et al., 1997; DE AZEVEDO et al., 2002).

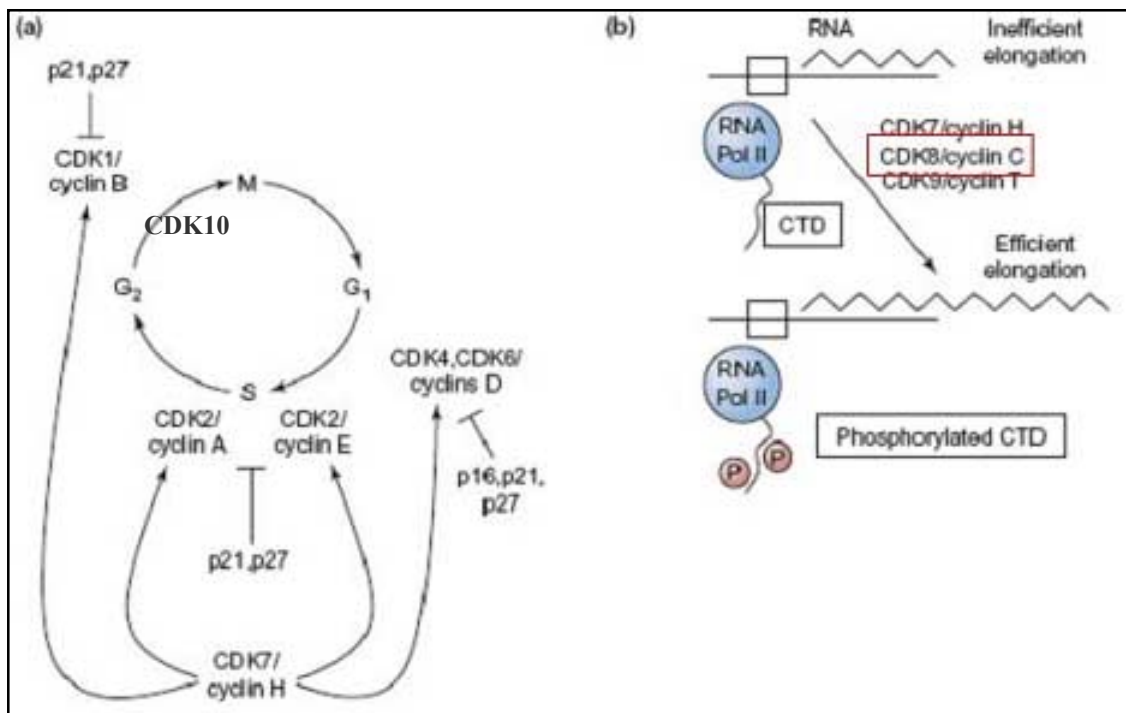


**Figura 1-8** Roscovitina, (MEIJER et al., 2002; SIELECKI et al., 2000).

### 1.9. CDKs 8 e 10 como alvos para desenho de inibidores

A CDK8 e a ciclina C são componentes do complexo da RNA polimerase II, onde funcionam como uma proteína quinase que fosforila o domínio C-terminal da subunidade maior da RNA polimerase II. O complexo protéico CDK8/ciclina C são encontrados em diversos mamíferos atuando como complexos protéicos mediadores, que reprimem a atividade de transcrição independentemente do domínio C-terminal *in vitro*. Segundo Akoulitchiev et al., 2000, o complexo CDK8/ciclina C pode regular a transcrição afetando as subunidades CDK7/ciclina H do fator de iniciação da transcrição, o fator IIH (TFIIH). A CDK8 fosforila a ciclina H, de mamíferos, na serina 5 e 304, ambos *in vivo* e *in vitro*, nas proximidades de seus domínios funcionais N- e C-terminal. Essa fosforilação reprime, tanto a atividade do TFIIH em ativar a transcrição, como a atividade quinase C-terminal. Portanto, a fosforilação da ciclina H pela CDK8, *in vivo*, apresenta um efeito negativo no crescimento celular.

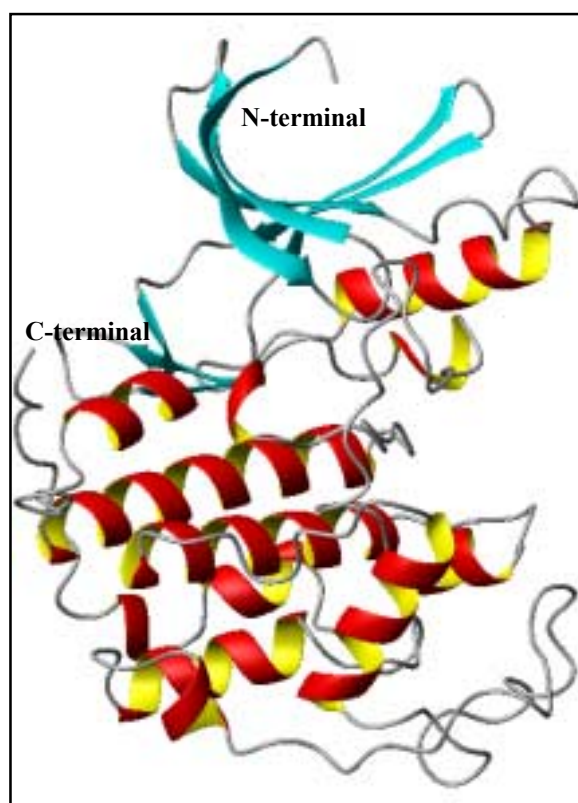
A CDK10 está envolvida no ciclo celular, na transição da fase G<sub>2</sub>/M. Ela interage com o N-terminal do fator de transcrição Ets2, que contém um domínio altamente conservado e um domínio de transativação.



**Figura 1-9** (a) A CDK10 está envolvida no ciclo celular, na transição da fase G<sub>2</sub>/M. (adaptado de SAUSVILLE, 2002). (b) A CDK8 e a ciclina C são componentes do complexo da RNA polimerase II, funcionam como uma proteína quinase que fosforila o domínio C-terminal da subunidade maior da RNA polimerase II.

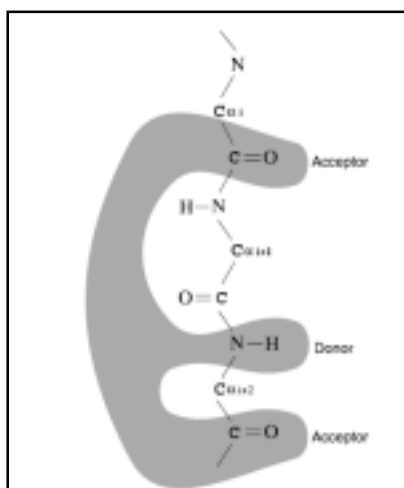
### 1.10. Um pouco sobre a CDK2 (molde) utilizado nas modelagens das CDKs 8 e 10

A CDK2 participa da regulação do ciclo celular, na fase G<sub>1</sub>, forma complexo com a ciclina D e na fase G<sub>2</sub>, forma complexo com a ciclina A. É uma proteína formada por 298 resíduos de aminoácidos, e sua massa molecular é de 34 kDa. No pequeno domínio N-terminal existem cinco fitas  $\beta$  antiparalelas que formam uma folha  $\beta$ , e também aparece uma única hélice  $\alpha$ . No domínio maior C-terminal, existem predominantemente hélices  $\alpha$  e há presença de uma pequena folha  $\beta$ , e um empacotamento *pseudo-4-helical*. Por isso a estrutura da CDK2 é chamada bilobal, devido à presença de dois domínios distintos, e entre esses domínios, está localizado o sítio de ligação do ATP, (CANDURI & DE AZEVEDO, 2005).



**Figura 1.10.** Estrutura da CDK2 (PDB:1HCL)

Extensivos estudos por cristalografia de difração de raios X foram realizados com a CDK2, e também com diversos inibidores. Com esses estudos chegou-se a conclusão que há um padrão de interação que se conserva, e, parece ser determinante no reconhecimento de pequenas moléculas pelo sítio de ligação de ATP para essa classe de enzimas. Esse padrão, denominado Garfo Molecular (DE AZEVEDO, et al., 2002) é composto por três ligações que envolvem, no caso do complexo CDK2-inibidor, o grupo C=O do resíduo Glu81, e os grupos N-H e C=O do resíduo Leu83, formando ligações de hidrogênio com o ligante.



**Figura 1-11.** Diagrama do Garfo Molecular (DE AZEVEDO, 2002)

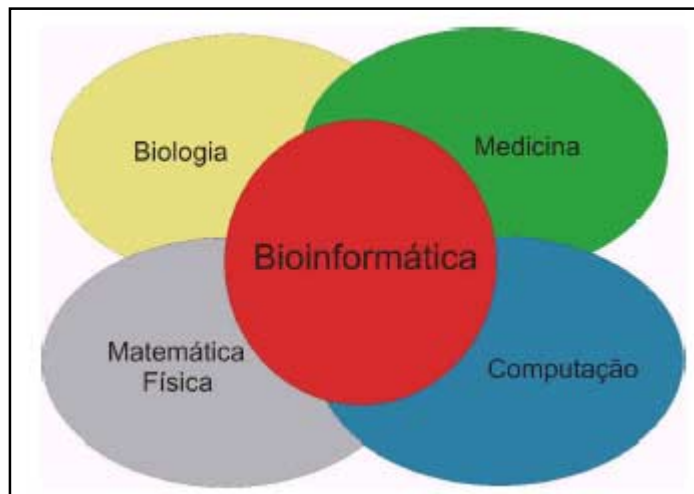
## **2. OBJETIVOS**

O objetivo deste trabalho é fornecer os modelos 3D das quinases dependentes de ciclinas 8 e 10 complexadas com inibidores específicos (Flavopiridol, Olomoucina e Roscovitina), bem como, com seu ligante natural ATP, uma vez que não possuem estruturas cristalográficas resolvidas. Estas CDKs são passíveis de estudo, pois estão codificadas no genoma humano. Para realização do processo de modelagem molecular há necessidade de um molde, que no caso, será a CDK2 com os respectivos ligantes: Flavopiridol, Olomoucina, Roscovitina, e ATP. A CDK2, complexada com esses inibidores específicos, foram cristalizadas por DE AZEVEDO, e ainda não foram depositadas no Banco de Dados de Proteínas. Espera-se que essas estruturas determinadas computacionalmente, produzam avanços no entendimento da função molecular e no mecanismo dessa família de proteínas, bem como, propor a compreensão da interação proteína/ligante, para possibilitar o desenho de drogas que inibam sua função.

### 3. METODOLOGIA E FERRAMENTAS

#### 3.1. Bioinformática

Uma definição ampla da bioinformática pode ser expressa da seguinte forma: é a aplicação de ferramentas de computação para análise, captura e interpretação de dados biológicos. Mostra-se uma área interdisciplinar e absorve a ciência da computação, matemática, biologia, física e medicina (Figura 3-1) (BAYAT, 2002).



**Figura 3-1** Interação das áreas que contribuem para a formação e o desenvolvimento da bioinformática (BAYAT, 2002).

Dentre os diversos objetivos da bioinformática, destaca-se o desenvolvimento e implementação de ferramentas que permitem o acesso, uso e manejo de vários tipos de informação, alinhamentos seqüenciais, estruturas, visualização de enovelamentos protéicos, etc. Essas novas ferramentas desenvolvidas através da bioinformática oferecem a oportunidade de interpretar maior quantidade de dados. A comparação seqüencial é a ferramenta computacional mais útil para os biólogos moleculares.

A crescente automação da biologia molecular e experimental e a aplicação da tecnologia da informação nas ciências biológicas conduzem a uma mudança fundamental na maneira como a pesquisa biológica é realizada (ARCURI et al, 2005).

### 3.2. Modelagem Molecular por Homologia ou Comparativa

A caracterização funcional de uma proteína é um dos problemas mais freqüentes na biologia (UCHÔA et al., 2004). A seqüência de uma proteína determina sua estrutura tridimensional, contudo o algoritmo que permita, com precisão, determinar essa estrutura, partindo de sua seqüência ainda está por ser determinado.

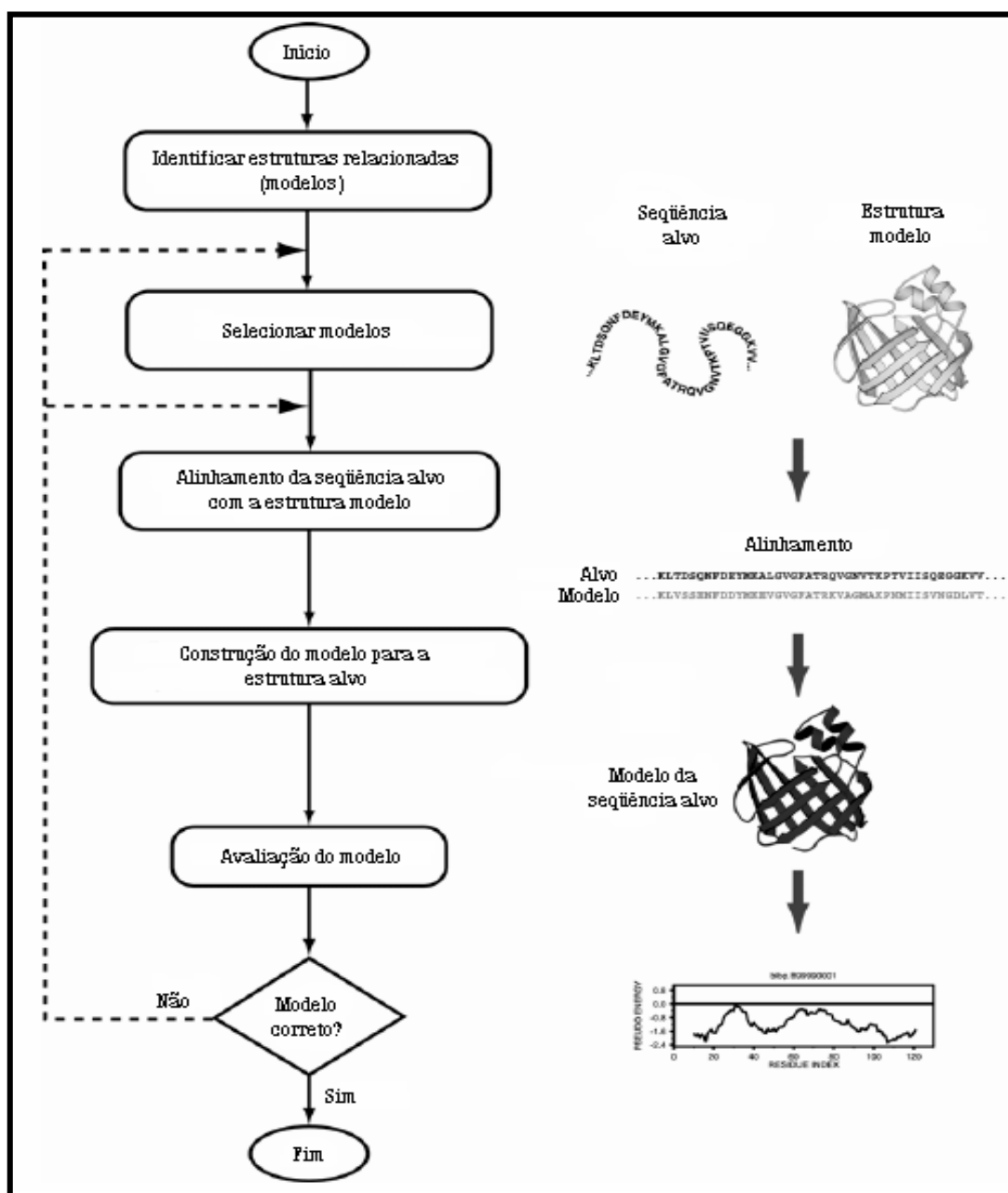
O método da modelagem molecular por homologia ou comparativa (BLUNDELL et al., 1987; MARTI-RENO et al., 2000), confia na similaridade detectável que mede a maioria das seqüências modeladas, ao menos, de uma estrutura conhecida.

Relacionamentos evolucionários permitem o uso de métodos computacionais para prever a estrutura de uma proteína baseada na sua similaridade com estruturas de proteínas conhecidas (ARCURI et al., 2004; PEREIRA et al., 2003; DA SILVEIRA et al., 2004). Entre todas as aproximações teóricas atuais, a modelagem molecular comparativa de estruturas de proteína é o único método computacional que tem o potencial de construir um modelo tridimensional para uma proteína de estrutura desconhecida (o alvo) baseada em uma ou mais proteínas relacionadas de estruturas conhecidas (moldes) com precisão e confiança comparáveis a estruturas experimentais de baixa resolução a partir da seqüência de aminoácidos (TRAMONTANO et al., 2001; BLUNDELL et al., 1987; GREER, 1981; JOHNSON et al., 1994; SALI & BLUNDELL, 1993; SALI, 1995; SANCHEZ & SALI, 1997; MARTI-RENO et al., 2000; FISER et al., 2002; SANCHEZ & SALI, 2000).

A precisão da modelagem comparativa é correlacionada com a identidade da seqüência, da porcentagem em que o modelo é baseado, imitando a correlação entre o estrutural e as similaridades da seqüência de duas proteínas. Erros incluem enganos na modelagem das cadeias laterais, distorções e deslocamentos em regiões de alinhamento, erros nas regiões sem molde, erros no alinhamento e moldes incorretos. Os erros de alinhamento aumentam rapidamente abaixo da identidade da seqüência de 30% e transformam-se na origem mais significativa dos erros em modelos comparativos. Os modelos comparativos têm sido usados em estudos de mecanismos catalíticos de enzimas, desenho e melhoramento de ligantes, *docking* de macromoléculas, *screening* virtuais, *docking* de pequenos ligantes, etc (FISER et al., 2002).

A utilidade da modelagem comparativa está aumentando firmemente porque o número de diferentes enovelamentos estruturais que as proteínas adotam é limitado e porque o número de estruturas novas determinadas experimentalmente está aumentando rapidamente. Esta tendência foi acentuada recentemente pelos projetos genômicos estruturais que tem como objetivo determinar ao menos uma estrutura para a maioria das famílias de proteínas. Uma previsão para que o genoma estrutural consiga seu objetivo é de 5 a 10 anos, fazendo modelagem comparativa aplicável à maioria das seqüências de proteínas (ARCURI et al, 2005). A modelagem comparativa representa um ganho de tempo na obtenção dessas estruturas 3D relacionadas, já que é independente das técnicas experimentais, como cristalografia e RMN (FISER et al., 2002).

A modelagem comparativa consiste em quatro etapas principais (SANCHEZ & SALI, 1998; ESWAR et al., 2003): (1) identificação da similaridade entre a seqüência alvo e ao menos uma estrutura de proteína conhecida (o molde); (2) alinhamento da seqüência alvo e a seqüência dos molde(s); (3) construção de um modelo baseado nos molde(s) escolhido(s), já que mais de um pode ser usado; e (4) avaliação do modelo para verificar sua qualidade (UCHÔA et al., 2004), (Figura 3-2).



**Figura 3-2** - Passos da modelagem molecular comparativa (Sanchez & Šali, 1997).

A localização no espaço 3D do maior número possível de átomos da proteína (alvo) depende do alinhamento. A construção da parte interna da proteína (alvo) baseia-se na idéia de que a conformação da cadeia principal da estrutura do molde pode ser transferida para ela. Existem muitos métodos de modelagem dessas regiões conservadas. Os mais importantes são: a modelagem por corpo rígido (GREER, 1990; BLUNDELL et al., 1987; BROWNE et al., 1969), a modelagem pela combinação de segmentos (CLAESSENS et al., 1989; LEVITT, 1992) e a modelagem pela satisfação de restrições espaciais (SALI & BLUNDELL, 1993)

(FILHO & ALENCASTRO, 2003); esta última, foi utilizada para as modelagens das CDKs 8 e 10, complexadas com seus inibidores específicos.

### **3.2.1. Procura e Seleção do Molde**

A primeira etapa do método é a identificação de pelo menos uma proteína cuja estrutura tridimensional é conhecida e que servirá de molde para a determinação da estrutura da proteína alvo. Duas situações são, em geral, possíveis: (1) a família protéica a que pertence à proteína alvo é conhecida e (2) não se sabe a que família a proteína alvo pertence (FILHO & ALENCASTRO, 2003).

No primeiro caso, é suficiente selecionar a proteína (molde) diretamente do PDB (BERMAN, et al. 2000, WESTBROOK et al., 2002). No segundo caso, o método mais simples é procurar de forma sistemática um ou mais moldes adequados em um banco de dados de estruturas (FILHO & ALENCASTRO, 2003) de proteínas conhecidas (PDB) (<http://www.rcsb.org/pdb>), usando como parâmetro de entrada uma sequência de estrutura não determinada (alvo) para que esta seja comparada com as sequências mais homólogas de estruturas conhecidas (moldes).

No caso deste trabalho, os moldes utilizados (CDK2-Flavopiridol, CDK2-Olomoucina e CDK2-Roscovitina) na modelagem molecular comparativa, não foram retirados de um banco de dados de estruturas de proteínas conhecidas (PDB), por não estarem ainda depositados no mesmo. No caso da CDK2-ATP, foi utilizado a primeira situação, e o código PDB do modelo é 1HCK .

### **3.2.2. Alinhamento e Construção do Modelo**

Uma vez feita a seleção do molde, um método de alinhamento deve ser usado para realizar um segundo alinhamento entre o molde com a sequência designada. O objetivo é alinhar resíduos estruturalmente equivalentes (FILHO & ALENCASTRO, 2003). Após o alinhamento, é possível reconhecer regiões estruturalmente conservadas e regiões variáveis (FILHO & ALENCASTRO, 2003). Para este trabalho, foi usado o programa MULTALIN (COPERT, 1998). Este alinhamento será utilizado pelo programa MODELLER para gerar os modelos (SALI & BLUNDELL, 1993).

Neste trabalho de modelagem das CDKs 8 e 10 complexadas com inibidores específicos, foi utilizado uma aproximação automatizada para modelagem comparativa de estrutura de proteína baseada na satisfação da restrição espacial (SALI & BLUNDELL, 1993; FISER, 2000) que é executada pelo programa MODELLER (<http://guitar.rockefeller.edu>) (SALI & BLUNDELL, 1993; FISER et al., 2002), através do servidor de *web* PARMODEL (UCHÔA et al., 2004).

A saída é um modelo 3D para a sequência alvo contendo todas as cadeias principais e cadeias laterais dos átomos sem o hidrogênio. No caso deste trabalho foram gerados 1000 modelos para cada complexo, e, a escolha dos melhores modelos foi baseada na qualidade estereoquímica.

### 3.2.3. Avaliação do Modelo

Diversas metodologias existem para a checagem e avaliação da qualidade de uma estrutura, seja ela determinada experimentalmente ou por modelagem. O modelo pode ser avaliado como um todo ou por regiões. Como a qualidade do modelo depende muito da estrutura da proteína (molde) utilizada, é preciso que esta última tenha sido obtida em alta resolução ( $\leq 2\text{\AA}$ ) e com um fator-R satisfatório ( $\leq 20\%$ ). A qualidade do modelo construído será proporcional à qualidade da estrutura 3D utilizada como molde (SANCHEZ & ŠALI, 1998).

A análise geométrica das estruturas de proteínas possui dois objetivos principais: (1) avaliação interna de sua consistência intrínseca que verifica se o modelo satisfaz ou não as restrições calculadas e (2) avaliação externa que confia em informações que não foram usadas no cálculo do modelo. A análise geométrica pode mostrar onde um modelo determinado experimentalmente ou por modelagem viola as leis da química. Existem leis da física que governam interações intermoleculares. Essas interações limitam não somente os contatos entre pares de átomos em diferentes partes de uma cadeia protéica, como também o tanto que os átomos possam girar livremente ao redor das ligações que os conectam. A estrutura das órbitas atômicas e a natureza das ligações entre átomos impõem limites naturais na posição de átomos ligados entre si, de modo que os ângulos de ligação e os ângulos diédricos fiquem, restritos a um grupo limitado de valores. Algumas são executadas de forma independente; outras são incorporadas em pacotes de simulação e de análise de estrutura mais abrangentes (GIBAS & JAMBECK, 2001).

A qualidade dos modelos das quinases dependentes de ciclinas 8 e 10 foram avaliadas através dos programas: WHATCHECK/WHATIF (HOOFT et al., 1996; VRIEND, 1990), VERIFY-3D (BOWIE et al., 1991; LUTHY et al., 1992; KABSCH & SANDER, 1983), RMSD da geometria ideal (ENGH & HUBER, 1991; KEARSLEY, 1989) usando o programa X-PLOR (SCHWIETERS et al., 2003; BRÜNGER, 1992), o RMSD entre a distância dos átomos de Ca - Ca sobrepostos (KABASH & SANDER, 1983) usando o programa LSQKAB do pacote CCP4 (Collaborative Computational Project N° 4., 1994), o G-FACTOR usando o programa PROCHECK (LASKOWSKI et al., 1994) e o diagrama de Ramachandran (RAMACHANDRAN et al., 1963; RAMAKRISHNAN & RAMACHANDRAN, 1965) usando o programa PROCHECK (LASKOWSKI et al., 1994).

Na avaliação da interação entre proteína-ligante foram verificados os potenciais eletrostáticos de superfícies gerados pelo PYMOL (DELANO, 2004).

Para obter a análise do programa VERIFY-3D (BOWIE et al., 1991; LUTHY et al., 1992; KABSCH & SANDER, 1983), e o diagrama de Ramachandran foi utilizado o módulo de avaliação do programa PARMODEL (UCHÔA et al., 2004). O módulo de avaliação PARMODEL permite que o usuário avalie os modelos gerados por modelagem, difração de raios X e RMN. O módulo requer como entrada apenas um arquivo das coordenadas atômicas no formato “pdb” da proteína que será avaliada, os resultados são mostrados na mesma página da web. As saídas geradas podem ser salvas em arquivos com extensão “pdf” para o PROCHECK, e para o VERIFY-3D, uma página de web com os valores e um gráfico.

### **3.3. Métodos de Avaliação**

- **PROCHECK**

O PROCHECK (LASKOWSKI et al., 1994) analisa a geometria global da estrutura ou cada resíduo individualmente, utilizando para isso parâmetros estereoquímicos derivados de estruturas de alta resolução ou bem refinadas (MORRIS et al., 1992) que constituem sua base de dados. É um pacote popular de programas para verificação da qualidade da estrutura 3D da proteína (ARCURI et al., 2005).

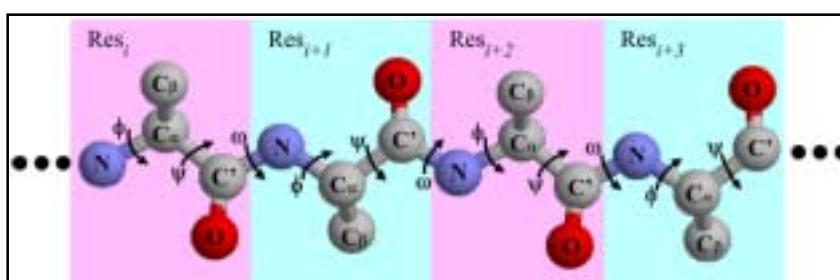
Os parâmetros estereoquímicos usados são aqueles descritos em detalhe em MORRIS et al. (1992). As verificações também fazem uso do comprimento e ângulo de ligação ideal, derivados de uma análise detalhada (ENGH & HUBER, 1991) de estruturas de pequenas

moléculas em Cambridge Structural Database (CSD) (ALLEN et al., 1979). Estes parâmetros utilizados como informações estereoquímicas checados pelo PROCHECK são: ligações covalentes, planaridade de grupos planares (aromáticos, ligações peptídicas, etc.) ângulos diédricos, quiralidade, interações não covalentes, ligações de hidrogênio da cadeia principal e pontes de dissulfeto.

O PROCHECK requer como entrada um arquivo das coordenadas atômicas da estrutura da proteína a ser avaliada no formato “pdb” e produz representações coloridas em PostScript (Adobe System Inc., 1985) facilmente interpretadas. Dentre as várias análises fornecidas pelo PROCHECK foram utilizadas neste trabalho: o Diagrama de Ramachandran e o G-Factor (ARCURI et al, 2005).

- **Diagrama de Ramachandran**

G. N. Ramachandran descreveu as conformações disponíveis para os aminoácidos em uma cadeia polipeptídica. A conformação da cadeia peptídica é simplesmente descrita pelos valores dos ângulos diédricos na estrutura principal da proteína (ângulo descrito pelo nitrogênio e carbono alfa ( $N-C_\alpha$ ) e o ângulo descrito pelo carbono alfa e carbono  $C_\alpha-C$ ) (Figura 3-3). Esses ângulos são denominados ângulos de torção  $\phi$  e  $\psi$ , respectivamente (GIBAS & JAMBECK, 2001).

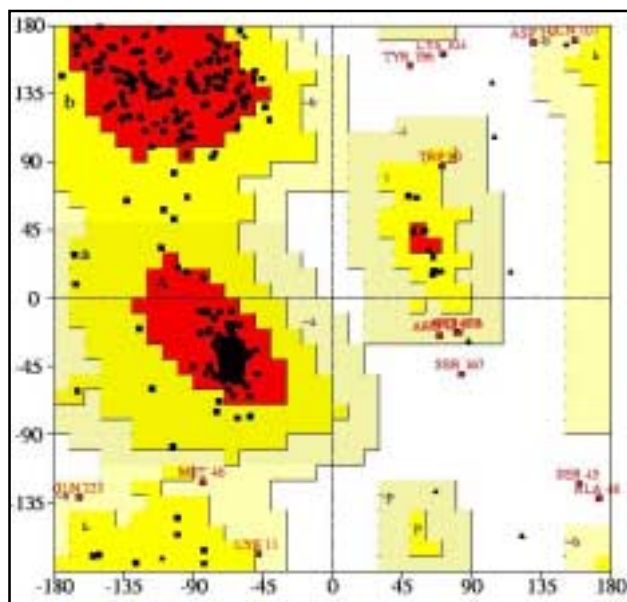


**Figura 3-3** Parte de uma cadeia polipeptídica indicando os graus de liberdade torcional de cada unidade peptídica e os ângulos de rotação  $\phi$  e  $\psi$ .

A cadeia não é livre para girar ao redor do terceiro tipo de ligação na estrutura principal da proteína, a ligação peptídica, por ser uma ligação parcialmente dupla e, portanto quimicamente restrita a ser planar; sendo assim, os valores de  $\phi$  e  $\psi$  para cada aminoácido fornecem uma descrição completa da estrutura da cadeia principal da proteína.

O diagrama de Ramachandran (RAMACHANDRAN et al., 1963; RAMAKRISHNAN & RAMACHANDRAN, 1965) indica a distribuição dos ângulos  $\Phi$  e  $\psi$  dos resíduos pertencentes a uma determinada estrutura. Seus indicadores de qualidade baseiam na análise de 118 estruturas de resolução de pelo menos 2,0 Å de resolução e  $R_{\text{factor}}$  menor que 20%. Nos seus critérios uma estrutura de boa qualidade deve ter 90% ou mais de seus resíduos nas regiões mais favoráveis (Figura 3-4). No diagrama abaixo, as áreas em branco correspondem a regiões onde existem os choques estereoquímicos na proteína. Estas regiões não permitidas estericamente são para todos os aminoácidos exceto a glicina (Gly), e a prolina (Pro), que é a única que não tem cadeia lateral e pode adotar os ângulos de torção  $\phi$  e  $\psi$  em todos os quadrantes do diagrama de Ramachandran.

As regiões vermelhas correspondem à conformação onde não há nenhum choque estereoquímico, ou seja, são as regiões permitidas e aonde se encontram as hélices- $\alpha$  e folhas- $\beta$ . As áreas em amarelo mostram as regiões limites onde um raio ligeiramente mais curto de van der Waals é usado no cálculo. Isto faz com que tenha uma região adicional que corresponda a hélices esquerda e direita.



**Figura 3-4** Diagrama de Ramachandran. A região generosamente permitida em vermelho, a região adicionalmente permitida em amarelo e a não permitida em branco. E as regiões onde estão as hélices- $\alpha$  (a e l) e folhas- $\beta$  (b).

Os aminoácidos L não podem formar extensas regiões de hélices esquerdas, mas ocasionalmente os resíduos individualmente podem adotar esta conformação. Estes resíduos são geralmente glicinas (Gly), mas também podem ser resíduos de asparagina ou aspartato onde a cadeia lateral forma uma ligação de hidrogênio com a cadeia principal e estabiliza conseqüentemente esta conformação inicialmente desfavorável. A hélice 3(10) ocorre perto da região hélice- $\alpha$  na parte superior a direita e está na borda da região adicionalmente permitida indicando uma estabilidade mais baixa. As regiões não permitidas envolvem geralmente um obstáculo estérico entre a cadeia lateral do grupo metileno  $C_\beta$  e átomos da cadeia principal.

- **G-FACTOR**

O G-factor fornece uma medida de uma propriedade estereoquímica, se é dada como normal ou não. Os valores gerais são obtidos a partir de uma média de todos os G-factors diferentes para cada resíduo da estrutura (LASKOWSKI et al., 1993). Valores de G-factors são computados para as seguintes propriedades (LASKOWSKI et al., 1993):

- 1) Ângulos de Torção: combinações de phi-psi; combinações de  $\chi_1$ - $\chi_2$ ; torções de  $\chi_1$  para aqueles resíduos que não possuem um  $\chi_2$ ; Ângulos de Torção combinados  $\chi_3$  e  $\chi_4$  e Ângulos de Torção Omega.
- 2) Geometria Covalente: calcula comprimentos de ligações entre cadeias principais e Ângulos de ligações entre cadeias principais.

- **VERIFY-3D**

O VERIFY-3D (BOWIE et al., 1991; LUTHY et al., 1992; KABSCH & SANDER, 1983) mede a compatibilidade entre a seqüência de aminoácidos de uma proteína e o modelo tridimensional, usando um perfil 3D. Para tanto, o programa utiliza a seguinte metodologia: 1) reduz a estrutura 3D a uma seqüência de ambientes químicos (um para cada resíduo) categorizados entre 18 possibilidades de acordo com a área do resíduo exposta ao solvente, com a fração de contatos feitos com átomos polares e a estrutura secundária local (perfil 3D), e 2) utilizando uma matriz (chamada matriz 3D-1D (BOWIE et al., 1991)) que descreve a probabilidade de se encontrar cada um dos 20 aminoácidos em cada uma das 18 classes de

ambientes químicos, o programa determina a probabilidade, por resíduo, de acordo com o tipo de aminoácidos e a natureza do ambiente químico calculado no passo anterior. A matriz 3D-1D é previamente conhecida e derivada de uma análise de estruturas conhecidas de boa qualidade (BOWIE et al., 1991; LUTHY et al., 1992).

O resultado é uma medida da compatibilidade entre a sequência e sua estrutura 3D descrita pelo seu perfil tridimensional. Lüthy et al. (1992) determinaram empiricamente o índice global esperado de  $S_{\text{calc}} = \exp(-0,83 + 1,008 \times \ln(L))$  de compatibilidade entre a sequência e a estrutura 3D onde L é o comprimento da sequência e  $S_{\text{calc}}$  representando, então, a soma das probabilidades individuais dos resíduos. Lüthy et al. (1992) também sugere um limite de  $0,45 \times S_{\text{calc}}$  para a confiabilidade da compatibilidade sequência/estrutura. Valores menores que  $0,45 \times S_{\text{calc}}$  indicam uma estrutura incorreta, valores em torno de  $0,45 \times S_{\text{calc}}$  podem estar corretas, mas possuem qualidade questionável e valores acima de  $0,45 \times S_{\text{calc}}$  indicam uma estrutura correta. O VERIFY-3D requer como entrada um arquivo de coordenadas atômicas no formato .pdb.

- **WHATIF/WHATCHECK**

O WHATIF (VRIEND, 1990) pode calcular quase todas as propriedades de proteínas, desde a área de superfície acessível ao solvente até valores de pKa para contato com a dinâmica molecular com GROMOS (VAN GUNSTEREN et al., 1987).

O WHATCHECK (HOOFT et al., 1996) fornece acesso a um subconjunto de verificações de qualidade estrutural WHATIF. WHATCHECK dá informações sobre a formação de regiões centrais hidrofóbicas, a acessibilidade de resíduos e átomos a moléculas de solvente (água), a distribuição espacial de grupos iônicos, a distribuição das distâncias atômicas e das ligações de hidrogênio da cadeia principal. Ou seja, retrata a estereoquímica, comprimentos de ligações, ângulos e diédricos, em forma de um relatório gerado no formato pdf.

- **RMSD (Desvio Médio Quadrático)**

O RMSD pode ser calculado de duas maneiras: (1) uma função de todos os átomos de uma proteína (ENGH & HUBER, 1991 – Fórmula 1) e (2) uma função de algum subconjunto

dos átomos, como a estrutura principal da proteína ou posições de  $C_{\alpha}$  apenas (KABASH & SANDER, 1983 – Fórmula 2).

$$rmsd_{C\alpha-C} = \sqrt{\sum_{j=1}^N (d_j - d_{C\alpha-C})^2} / N$$

**Fórmula (1)**

$$rmsd = \sqrt{\sum_{j=1}^N d_j^2} / N$$

**Fórmula (2)**

O RMSD da geometria ideal (KEARSLEY, 1989) é calculado como uma função de todos os átomos de uma proteína, podemos adotar a seguinte definição, que o RMSD da geometria ideal serve para analisar a medida da variação da posição de cada átomo de uma estrutura com relação aos parâmetros observados por ENGH e HUBER. O RMSD da geometria ideal utiliza como parâmetros o comprimento e ângulo de ligação ideal, derivados de uma análise detalhada (ENGH & HUBER, 1991) de estruturas de pequenas moléculas em *Cambridge Structural Database* (CSD) (ALLEN et al., 1979). O programa utilizado para executar os cálculos de comprimento e ângulos de ligação foi o programa X-PLOR (SCHWIETERS et al., 2003; BRÜNGER, 1992).

O RMSD da sobreposição é calculado como uma função utilizando um subconjunto de átomos da proteína. Pois o RMSD da sobreposição mede a diferença estrutural entre duas estruturas sobrepostas, normalmente o subconjunto de átomos escolhidos são os átomos da cadeia principal ( $C_{\alpha}$ -  $C_{\alpha}$ ). O RMSD é uma função da distância entre os átomos em uma estrutura e os mesmos átomos em outra estrutura.

Para calcular o RMSD da sobreposição as duas estruturas em consideração devem primeiro ser superpostas. A sobreposição das estruturas de proteínas começa geralmente com uma comparação de seqüências. A comparação de seqüências define as relações entre os pares de átomos onde o RMSD da sobreposição é calculado. Uma vez definida as relações átomo-a-átomo entre duas estruturas, com um programa de sobreposição como o LSQKAB do pacote CCP4 (Collaborative Computational Project N° 4., 1994), alcançamos uma superposição ótima entre as duas estruturas, isto é, a superposição com o menor RMSD possível. Os algoritmos de superposição otimizam a orientação e a posição espacial de duas moléculas (translação e rotação) entre si. Uma vez efetuadas as superposições ótimas de todos os pares de estruturas, obtemos o valor do RMSD da sobreposição (GIBAS & JAMBECK, 2001).

- **Visualização dos Modelos**

Para visualização dos elementos de estrutura secundária das proteínas modeladas foi usado o programa PYMOL (DELANO, 2004) e para a análise de interações proteína-ligante foi usado o programa XtalView (MCREE, 1999).

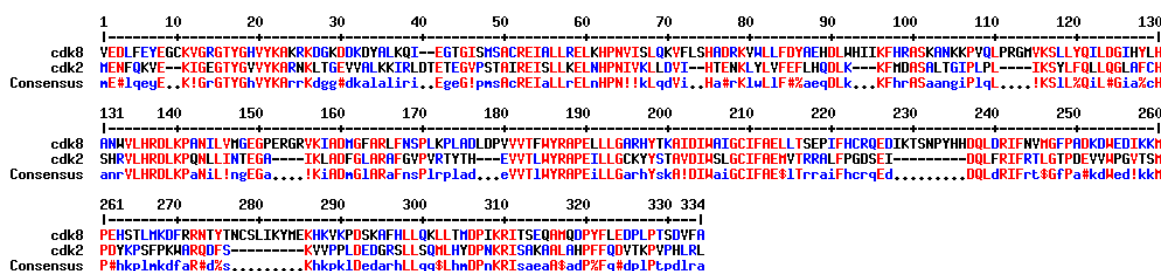
Essas ferramentas descritas, foram utilizadas para auxiliar nas análises quantitativas e qualitativas das proteínas em estudo, possibilitando a avaliação e a validação dos modelos gerados através do método de modelagem molecular comparativa.

## 4. ANÁLISE E RESULTADOS

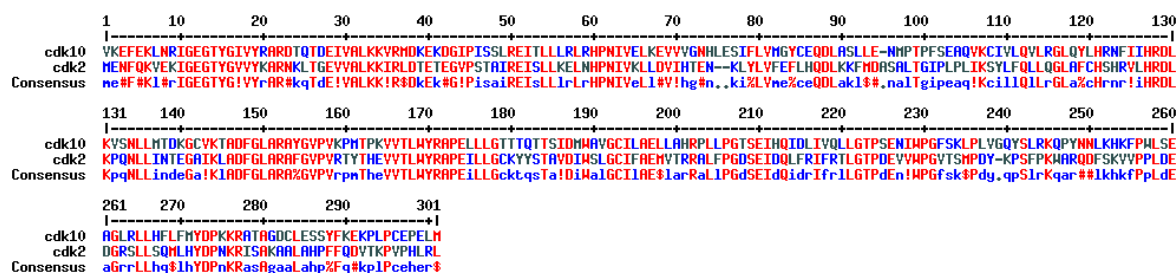
Foram construídos 4 modelos para CDK8 e 4 modelos para CDK10, na presença dos ligantes ATP, flavopiridol, olomoucina, e roscovitina. Em ambos, os casos, foram utilizados o mesmo molde na presença dos diferentes ligantes, CDK2 com ATP (PDB: 1HCK), com flavopiridol (PDB:não depositado), com olomoucina (PDB:não depositado), e com roscovitina (PDB:não depositado). Foram gerados 1000 modelos para cada complexo CDK-ligante, e a escolha do melhor modelo foi baseada na qualidade estereoquímica utilizando o programa PROCHECK (LASKOWSKI et al. ,1993), e a função objetiva do MODELLER (SALI & BLUNDELL,1993).

### 4.1. Comparação das Sequências Primárias

As porcentagens de identidade das seqüências primárias das CDK8 (SWISS-PROT:P49336) e CDK10 (SWISS-PROT:Q15131), comparadas com a seqüência primária da CDK2 (SWISS-PROT:P24941), indicam que essa estrutura é um bom molde a ser considerado para a modelagem de ambas, CDK8 e CDK10, pois apresentam identidade superior a 30%. A porcentagem de identidade da CDK8 em relação à CDK2 é de 35,5 %, e com a CDK10, a porcentagem é de 43 %. Os alinhamentos são mostrados na (Figura 4-1(a) e (b)), e foram executados pelo MULTALIN (COPERT, 1998). Na figura de alinhamento, os espaços (gaps) são representados por linhas tracejadas, os resíduos idênticos estão representados em vermelho, os similares em azul, e os resíduos não idênticos em preto.



(a)



(b)

**Figura 4-1** Alinhamento das seqüências primárias em (a) CDK8 e CDK2; em (b) CDK10 e CDK2; observando que os cortes necessários já foram efetuados nas terminações N e C-terminais, das CDK8 e CDK10.

## 4.2. Qualidade do Modelo

Avaliar a qualidade do modelo é um quesito fundamental, pois ela será proporcional à qualidade da estrutura 3D utilizada como molde (SANCHEZ et al., 1998). Por isso, deve-se avaliar o modelo, como um todo, e, por região. Os modelos foram construídos baseados nas coordenadas atômicas das estruturas resolvidas por cristalografia. As coordenadas atômicas de todas as moléculas de água foram retiradas dos moldes.

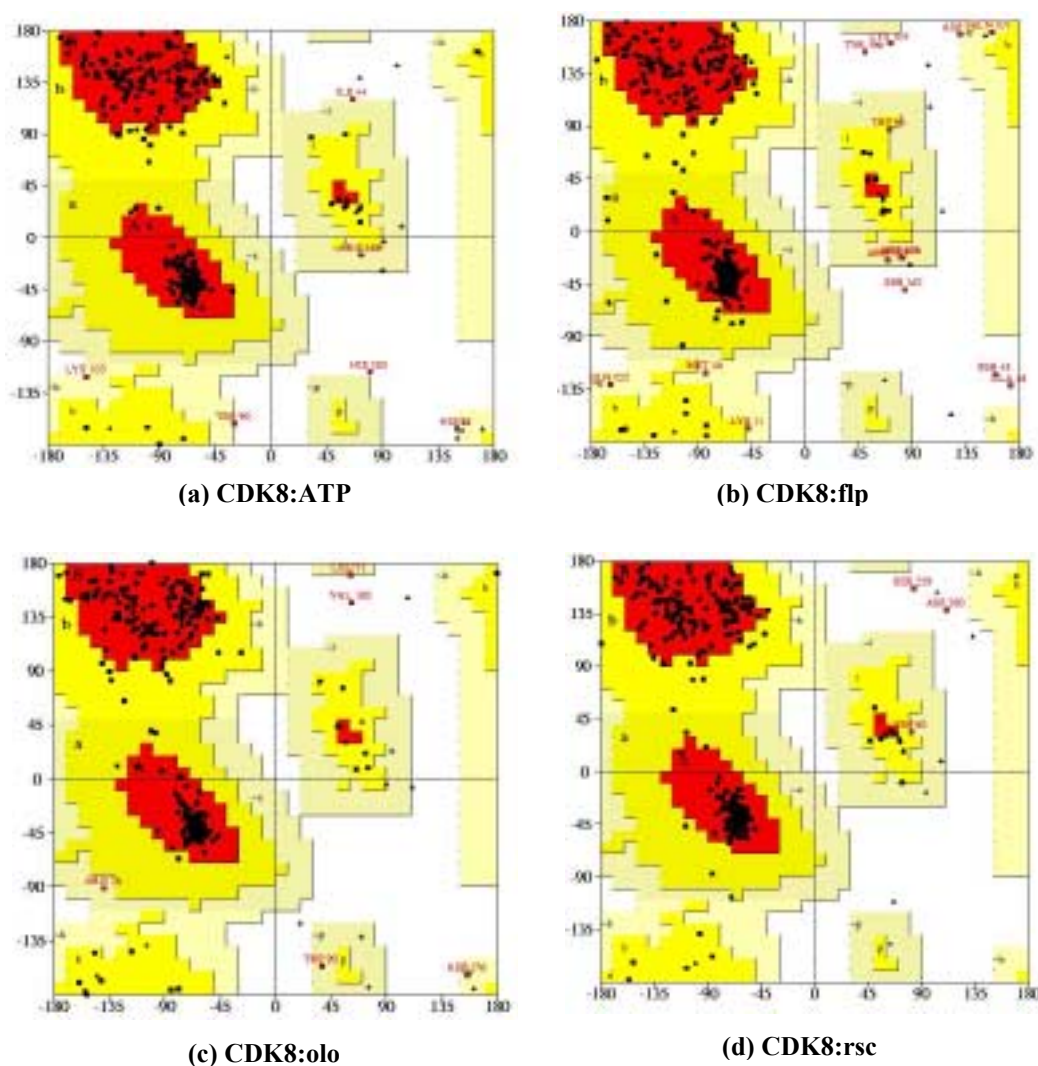
Após a etapa inicial da escolha do melhor modelo, análises mais criteriosas e detalhadas foram efetuadas com utilização do VERIFY 3D, WHATCHECK, RMSD da geometria ideal e PROCHECK, a fim de verificar a qualidade do modelo construído. Os resultados obtidos do modelo final foram comparados com o valor da estrutura 3D (molde).

O programa PROCHECK foi utilizado para analisar a geometria global dos modelos através do Diagrama de Ramachandran (Tabelas 4-1 e 4-2) e (Figuras 4-2 (a), (b), (c) e (d) e 4-3(a), (b), (c) e (d)). A análise dos diagramas de Ramachandran dos modelos da CDK8 (alvo), mostrou que o melhor modelo possui mais de 90,0% dos resíduos sobre a região permitida e para a CDK10 (alvo) mostrou que o melhor modelo possui mais de 93,0% dos resíduos sobre a região permitida.

**Tabela 4-1** Resultados das análises do Diagrama de Ramachandran dos modelos (CDK8 complexada com: ATP, flp,olo,rsc) e de seus moldes (CDK2 complexada com: ATP, flp,olo,rsc), que estão em parênteses.

| Complexo  | Regiões do Gráfico de Ramachandran (%) |                          |                         |           |
|-----------|----------------------------------------|--------------------------|-------------------------|-----------|
|           | Permitida                              | Adicionalmente Permitida | Generosamente Permitida | Proibida  |
| CDK8: ATP | 89,2 (88,6)                            | 8,8 (10,6)               | 1,0 (0,8)               | 1,0 (0,0) |
| CDK8: flp | 83,1 (79,7)                            | 12,2 (14,2)              | 2,7 (4,2)               | 2,0 (1,9) |
| CDK8: olo | 87,8 (87,9)                            | 10,5 (10,9)              | 1,4 (0,8)               | 0,3 (0,4) |
| CDK8: rsc | 90,9 (87,1)                            | 8,1 (12,9)               | 0,3 (0,0)               | 0,7 (0,0) |

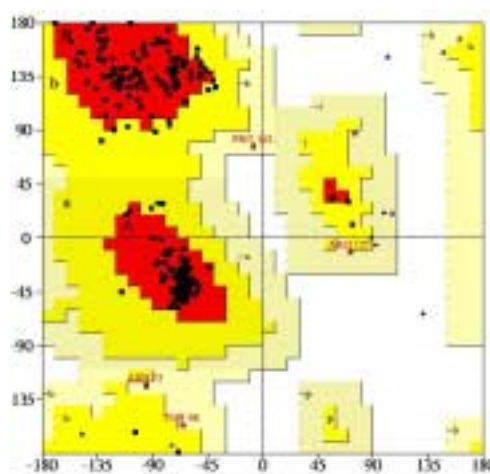
Nas figuras 4-2 (a) até (d) e 4-3 (a) até (d) estão os gráficos de Ramachandran. As regiões em vermelho correspondem às regiões permitidas, as regiões em amarelo escuro são as adicionalmente permitidas, em amarelo claro estão as regiões generosamente permitidas e as regiões em branco são as proibidas. Os resíduos de glicina e prolina, são representados por triângulos.



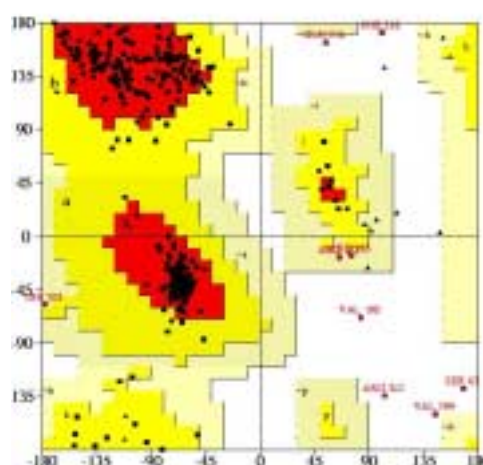
**Figura 4-2** Diagramas de Ramachandran das (a) CDK8: ATP, (b) CDK8: flp, (c) CDK8: olo, (d) CDK8: rsc .

**Tabela 4-2** Resultados das análises do Diagrama de Ramachandran dos modelos (CDK10 complexada com os ligantes: ATP, flp, olo, rsc) e de seus moldes (CDK2 complexada com os ligantes: ATP, flp, olo, rsc) , que estão

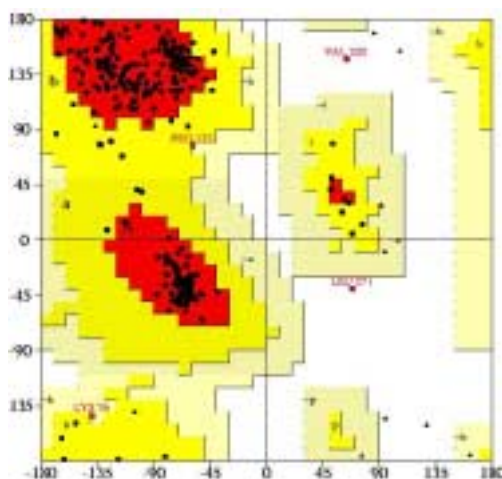
| Complexo   | Regiões do Gráfico de Ramachandran (%) |                          |                         |           |
|------------|----------------------------------------|--------------------------|-------------------------|-----------|
|            | Permitida                              | Adicionalmente Permitida | Generosamente Permitida | Proibida  |
| CDK10: ATP | 93,9 (88,6)                            | 5,0 (10,6)               | 1,1 (0,8)               | 0,0 (0,0) |
| CDK10: flp | 87,5 (79,7)                            | 9,8 (14,2)               | 1,2 (4,2)               | 1,6 (1,9) |
| CDK10: olo | 91,5 (87,9)                            | 7,3 (10,9)               | 0,4 (0,8)               | 0,8 (0,4) |
| CDK10: rsc | 92,2 (87,1)                            | 7,8 (12,9)               | 0,0 (0,0)               | 0,0 (0,0) |



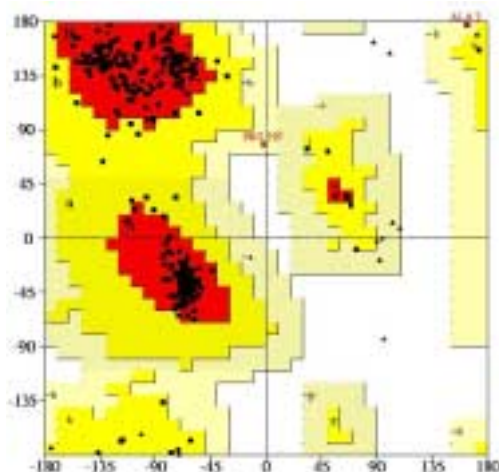
(a) CDK10:ATP



(b) CDK10:flp



(c) CDK10:olo



(d) CDK10:rsc

**Figura 4-3** Diagramas de Ramachandran das (a) CDK10: ATP, (b) CDK10: flp, (c) CDK10: olo, (d) CDK10: rsc .

Como se pode observar, os modelos apresentaram uma qualidade estereoquímica ligeiramente melhor do que a dos moldes utilizados.

Os resíduos que estão na região não permitida, estão localizados em uma região de superfície, ou seja, estão em contato com o meio (solvente); ou em uma região intermediária entre o solvente e o sítio da proteína, por isso, não comprometem a qualidade dos modelos. Resíduos de superfície (expostos ao solvente), geralmente, possuem maior flexibilidade de rotação, do que resíduos “enterrados” na proteína, não estão próximos o suficiente, da região de sítio ativo, para causarem choques estereoquímicos. Para analisar tais resíduos, foi utilizado o programa XtalView (McREE,1999).

Através do programa WHATCHECK foi constatado que os ambientes químicos de resíduos individuais se enquadram nos padrões freqüentemente encontrados em estruturas protéicas. Segundo os valores de referência do programa para modelagem molecular, o modelo final pode ser considerado de boa qualidade.

O G-factor global é uma medida da normalidade da estrutura como um todo. Representa uma média para o G-factor de cada resíduo na estrutura. O G-factor médio dos ângulos de torção e geometria covalente representa uma medida do desvio de uma dada propriedade estereoquímica. O valor ideal do G-factor deve ser acima de -0,5; no intervalo de -0,5 a -1,0 existem resíduos na proteína que deverão ser investigados e, abaixo de -1,0, os modelos apresentam qualidade ruim e deverão ser investigados (MORRIS et al., 1986). O resultado dos valores do G-factor, para os ângulos de torção, geometria covalente e para o G-factor global estão colocados na (Tabela 4-3) para os modelos de CDK8 e na (Tabela 4-4) para os modelos de CDK10:

**Tabela 4-3** Resultado dos valores de G-FACTOR dos modelos da CDK8 com ATP, flp, olo e rsc. Entre parênteses estão os valores obtidos para as estruturas 3D usadas como moldes.

| Modelos   | G-FACTOR          |                     |               |
|-----------|-------------------|---------------------|---------------|
|           | Ângulos de Torção | Geometria Covalente | Global        |
| CDK8: ATP | -0,12 ( 0,05)     | -0,40 ( 0,40)       | -0,21 ( 0,19) |
| CDK8: flp | -0,29 (-0,71)     | -0,56 (-0,73)       | -0,38 (-0,70) |
| CDK8: olo | -0,15 (-0,20)     | -0,60 (-1,14)       | -0,30 (-0,70) |
| CDK8: rsc | -0,14 (-0,48)     | -0,42 ( 0,45)       | -0,23 (-0,11) |

**Tabela 4-4** Resultado dos valores do G-factor dos modelos de CDK10 com ATP, flp, olo e rsc. Entre parênteses estão os valores obtidos para as estruturas 3D usadas como moldes.

| Modelos    | G-FACTOR          |                     |               |
|------------|-------------------|---------------------|---------------|
|            | Ângulos de Torção | Geometria Covalente | Global        |
| CDK10: ATP | -0,15 (-0,05)     | -0,93 ( 0,40)       | -0,42 ( 0,19) |
| CDK10: flp | -0,29 (-0,71)     | -0,42 (-0,73)       | -0,33 (-0,70) |
| CDK10: olo | -0,07 (-0,20)     | -0,37 (-1,14)       | -0,17 (-0,48) |
| CDK10: rsc | -0,17 (-0,48)     | -0,31 (0,45)        | -0,21 (-0,11) |

Para verificar a confiabilidade na compatibilidade das seqüências/estruturas foi utilizado o programa VERIFY 3D. O valor de Escore Total, o valor de Escore Ideal e o Escore  $S_{calc}$  final são mostrados nas (Tabela 4-5) e (Tabela 4-6). Os valores obtidos indicam que as estruturas dos modelos finais possuem compatibilidade entre suas seqüências primárias e as estruturas 3D construídas, pois os valores gerados para os modelos finais ficaram acima do limite de 0,45.

**Tabela 4-5** Resultado dos valores do VERIFY 3D dos modelos de CDK8 com seus respectivos complexos, os moldes de CDK2 estão entre parênteses

| Modelos   | VERIFY-3D       |                 |                   |
|-----------|-----------------|-----------------|-------------------|
|           | Escore Total    | Escore Ideal    | Escore $S_{calc}$ |
| CDK8: ATP | 117,27 (133,19) | 151,65 (134,16) | 0,77 (0,99)       |
| CDK8: flp | 116,11 (128,81) | 151,65 (130,48) | 0,77 (0,99)       |
| CDK8: olo | 116,25 (129,94) | 151,65 (134,62) | 0,77 (0,97)       |
| CDK8: rsc | 94,32 (109,38)  | 151,65 (136,00) | 0,62 (0,80)       |

**Tabela 4-6** Resultado dos valores do VERIFY 3D dos modelos de CDK10 com seus respectivos complexos, os moldes de CDK2 estão entre parênteses

| Modelos    | VERIFY-3D       |                 |                   |
|------------|-----------------|-----------------|-------------------|
|            | Escore Total    | Escore Ideal    | Escore $S_{calc}$ |
| CDK10: ATP | 107,25 (133,19) | 136,92 (134,16) | 0,78 (0,99)       |
| CDK10: flp | 108,13 (128,81) | 164,55 (130,48) | 0,66 (0,99)       |
| CDK10: olo | 108,77 (129,94) | 164,55 (134,62) | 0,66 (0,97)       |
| CDK10: rsc | 90,41 (109,38)  | 164,55 (136,00) | 0,55 (0,80)       |

O valor do RMSD da geometria ideal (Tabela 4-7) e (Tabela 4-8) para o comprimento e ângulo de ligação foi calculado pelo programa X-PLOR (SCHWIETERS et al., 2003; BRÜNGER, 1992), que gera um arquivo com os valores do RMSD para as ligações e ângulos diédricos e impróprios bem como contatos não ligados, os valores ideais para o comprimento e ângulo de ligação devem estar abaixo 0,1 Å e 10,0° respectivamente, para que a estrutura da proteína como um todo, fique validada.

**Tabela 4-7** Resultado dos valores do RMSD da geometria ideal dos modelos de CDK8.\*Entre parênteses estão os valores obtidos para as estruturas 3D usadas como moldes (CDK2).

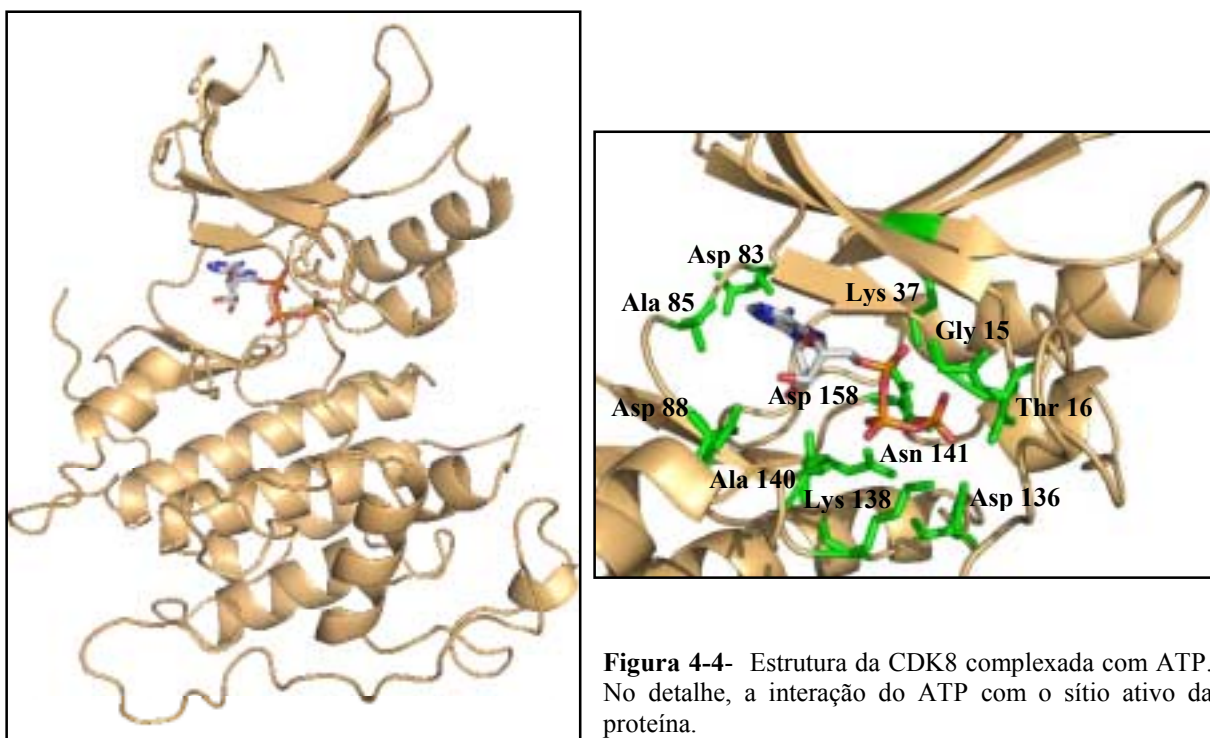
| Modelos   | RMSD Geometria Ideal       |                       |
|-----------|----------------------------|-----------------------|
|           | Comprimento de Ligação (Å) | Ângulo de Ligação (°) |
| CDK8: ATP | 0,020 (0,012)              | 2,461 (1,697)         |
| CDK8: flp | 0,020 (0,024)              | 2,637 (2,825)         |
| CDK8: olo | 0,020 (0,013)              | 2,607 (3,644)         |
| CDK8: rsc | 0,019 (0,012)              | 2,503 (1,586)         |

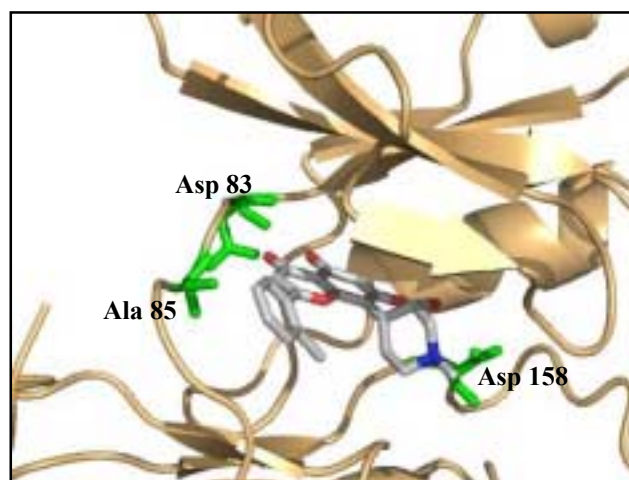
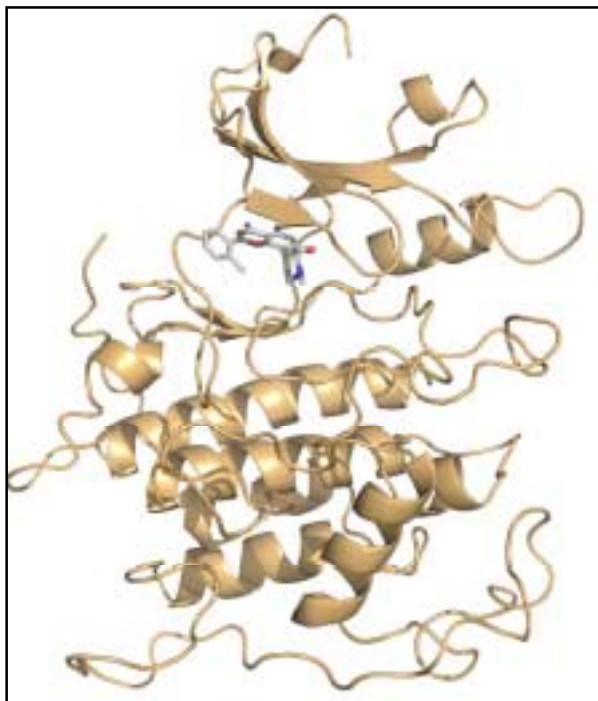
**Tabela 4-8** Resultado dos valores do RMSD da geometria ideal dos modelos de CDK10.\*Entre parênteses estão os valores obtidos para as estruturas 3D usadas como moldes (CDK2).

| Modelos    | RMSD Geometria Ideal       |                       |
|------------|----------------------------|-----------------------|
|            | Comprimento de Ligação (Å) | Ângulo de Ligação (°) |
| CDK10: ATP | 0,022 (0,012)              | 2,832 (1,697)         |
| CDK10: flp | 0,021 (0,024)              | 2,515 (2,825)         |
| CDK10: olo | 0,020 (0,013)              | 2,393 (3,644)         |
| CDK10: rsc | 0,020 (0,012)              | 2,388 (1,586)         |

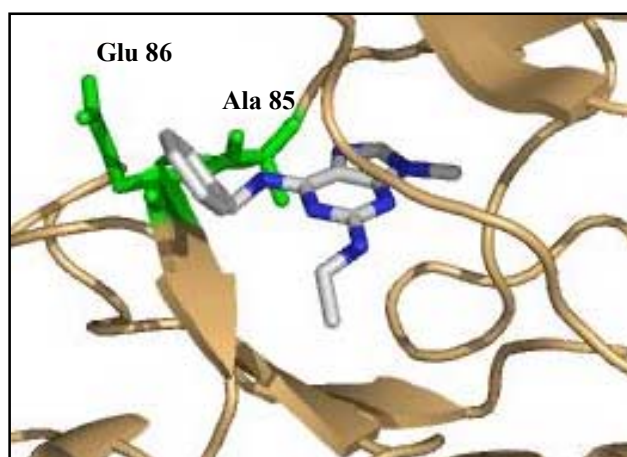
### 4.3. Descrição dos Modelos

Os modelos da CDK8 e CDK10 são similares à estrutura da CDK2. São compostas de dois lóbulos, um lóbulo N-terminal composto predominantemente de fitas  $\beta$  pregueadas e um C-terminal composto de hélices  $\alpha$ . Tratam-se de estruturas tipicamente bilobais. A característica das quinases é a presença do sítio catalítico localizado entre estes dois lóbulos. Devido a esta similaridade é possível observar as moléculas dos ligantes posicionadas entre os dois lóbulos, no sítio de ligação, como ocorre com a CDK2. Esses elementos de estruturas terciárias são mostrados nas (Figuras 4-4 a 4-11), gerados pelo programa PYMOL (DELANO, 2004).

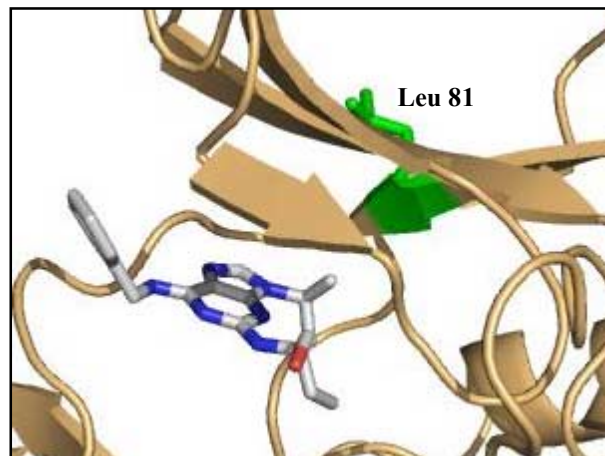
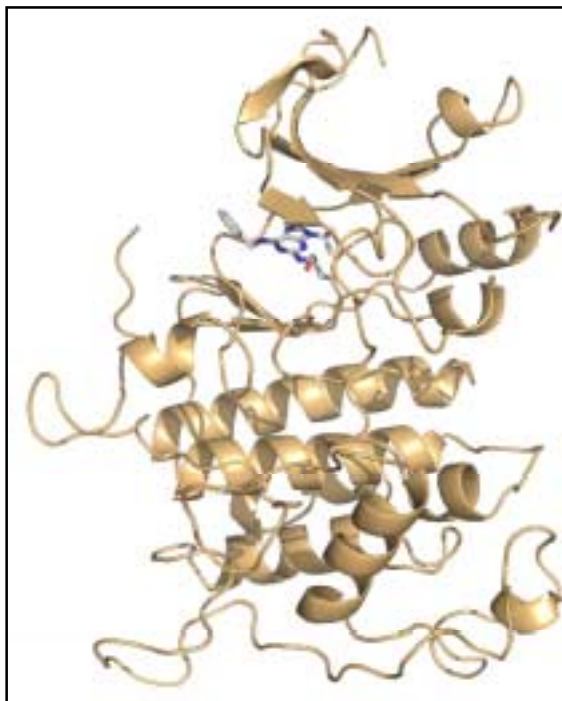




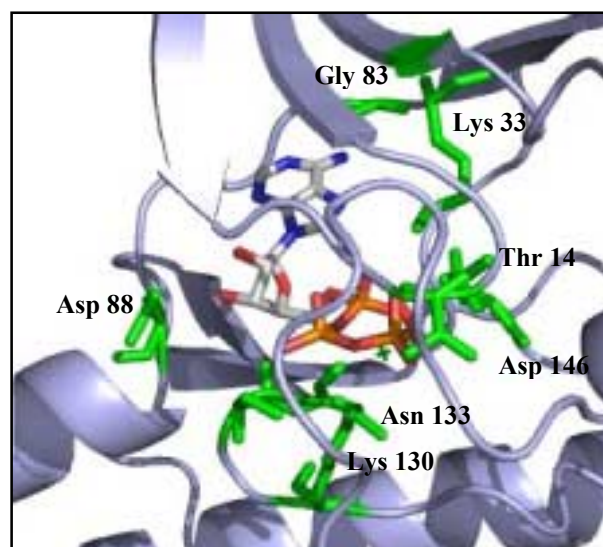
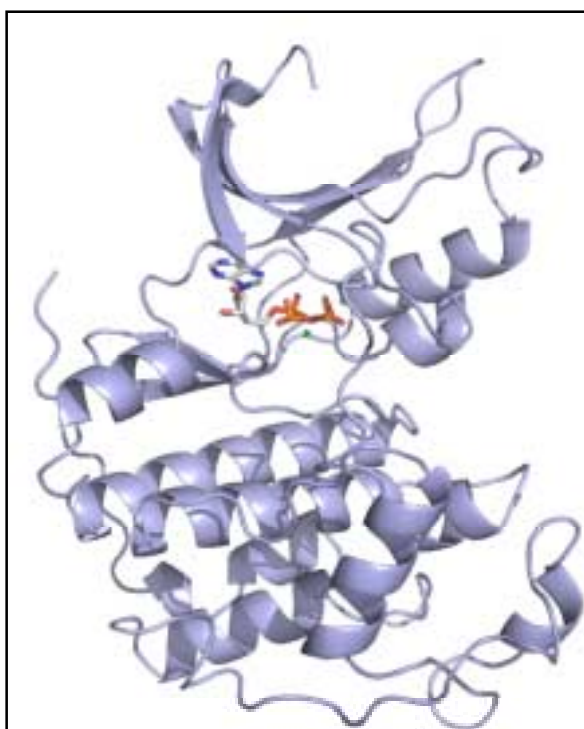
**Figura 4-5-** Estrutura da CDK8 complexada com flavopiridol. No detalhe, a interação do flavopiridol com o sítio ativo da proteína.



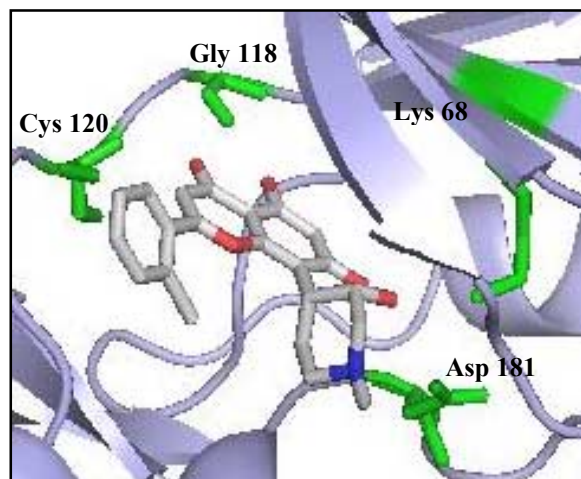
**Figura 4-6-** Estrutura da CDK8 complexada com olomoucina. No detalhe, a interação da olomoucina com o sítio ativo da proteína.



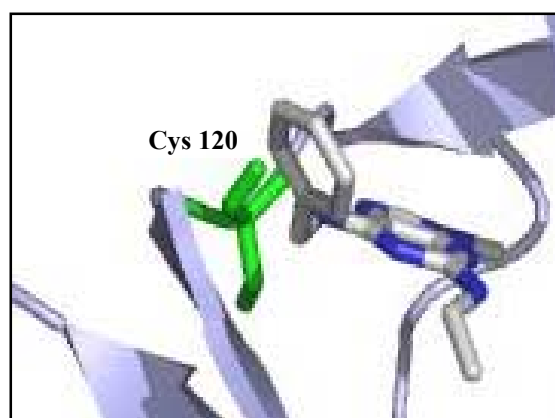
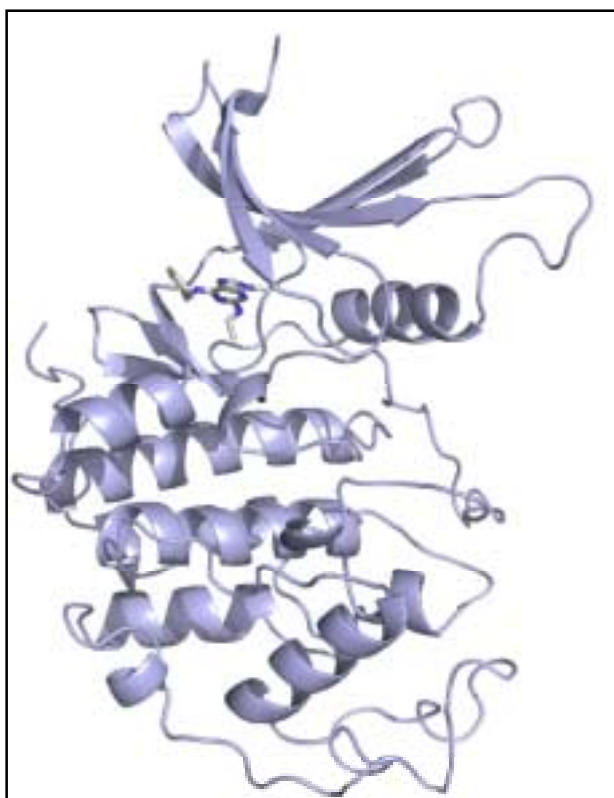
**Figura 4-7-** Estrutura da CDK8 complexada com roscovitina. No detalhe, a interação da roscovitina com o sítio ativo da proteína.



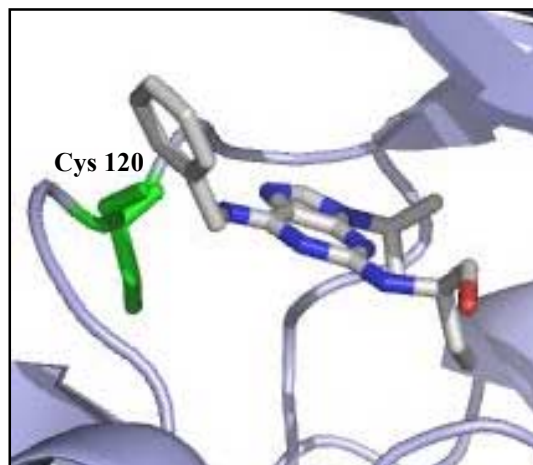
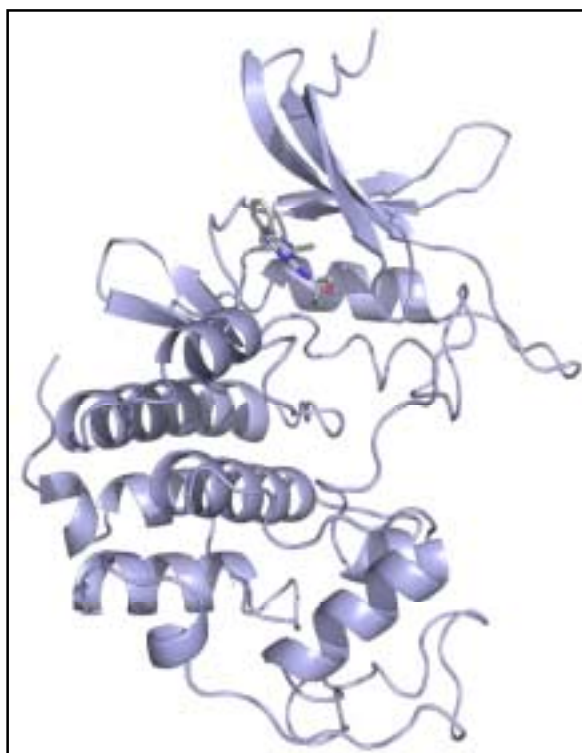
**Figura 4-8-**Estrutura da CDK10 complexada com ATP. No detalhe, a interação do ATP com o sítio ativo da proteína.



**Figura 4-9-** Estrutura da CDK10 complexada com flavopiridol. No detalhe, a interação da flavopiridol com o sítio ativo da proteína.



**Figura 4-10-** Estrutura secundária da CDK10 complexada com olomoucina. No detalhe, a interação da olomoucina com o sítio ativo da proteína.

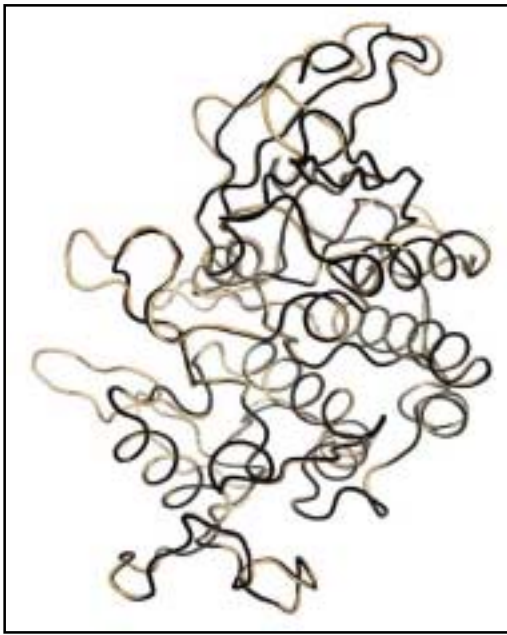


**Figura 4-11-** Estrutura da CDK10 complexada com roscovitina. No detalhe, a interação da roscovitina com o sítio ativo da proteína.

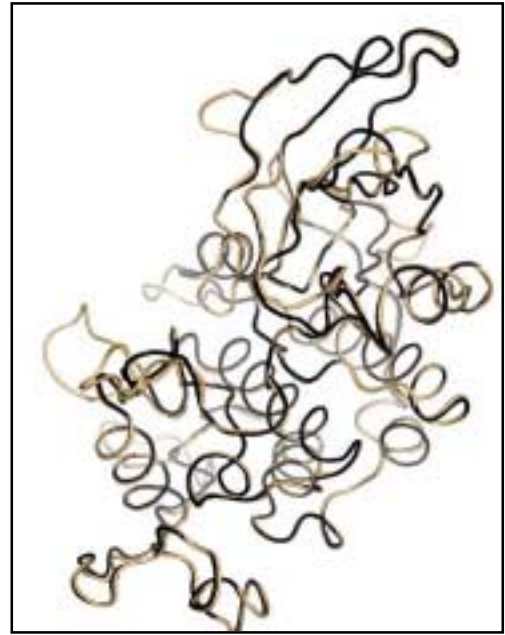
No cálculo do RMSD da sobreposição foi considerado somente  $C_{\alpha} - C_{\alpha}$  com o auxílio do programa LSQKAB do pacote CCP4 (Collaborative Computational Project N° 4., 1994). Com a sobreposição é possível verificar se há possíveis alterações no posicionamento das cadeias laterais dos resíduos ou se há alguma alteração na conformação da estrutura da proteína, além de determinar o RMSD da sobreposição (Tabela 4-9) e (Figura 4-12) e (Tabela 4-10) (Figura 4-13). A sobreposição é feita entre o modelo e o molde utilizando as coordenadas atômicas de cada um.

**Tabela 4-9** Valores de RMSD da sobreposição entre a CDK8 e a CDK2.

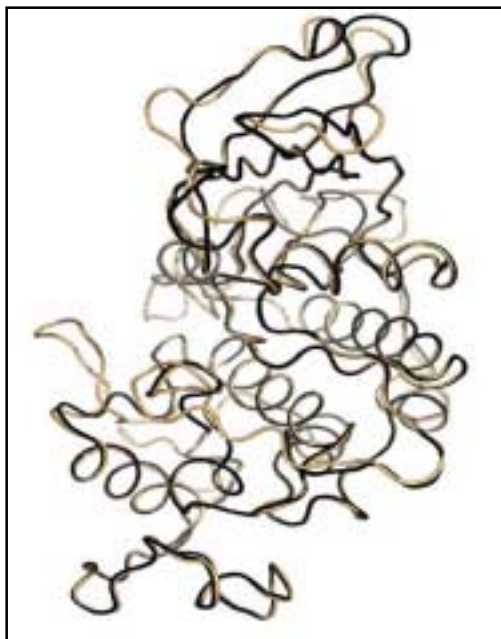
| MODELOS   | RMSD<br>sobreposição (Å) |
|-----------|--------------------------|
| CDK8: ATP | 0,119                    |
| CDK8: flp | 0,122                    |
| CDK8: olo | 0,185                    |
| CDK8: rsc | 0,146                    |



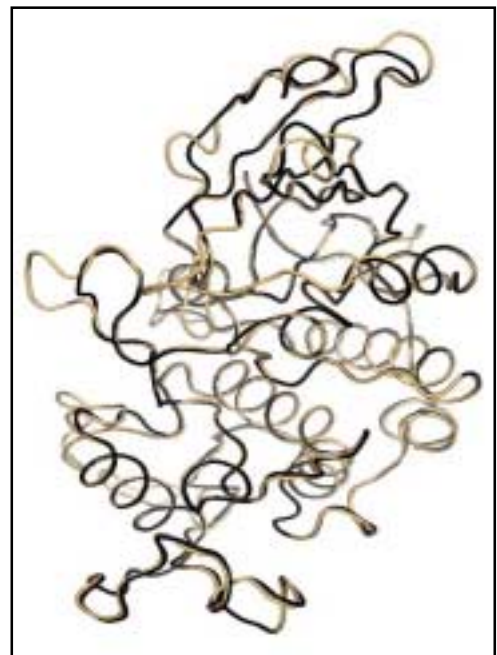
(a)



(b)



(c)

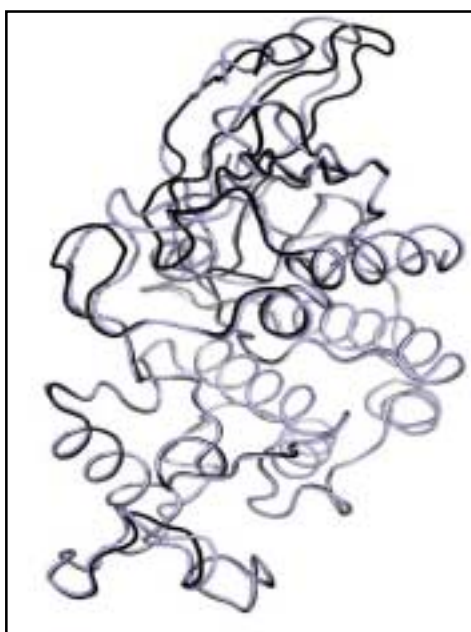


(d)

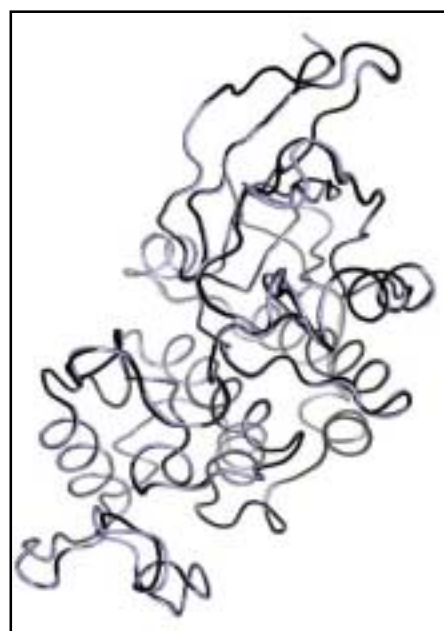
**Figura 4-12** Sobreposições da CDK8 (linha amarela) com seu molde CDK2 (linha preta); em (a) CDK8:ATP com CDK2:ATP, em (b) CDK8:flp com CDK2:flp, (c) CDK8:olo com CDK2:olo e em (d) CDK8:rsc com CDK2:rsc .

**Tabela 4-10** Valores de RMSD da sobreposição entre a CDK10 e a CDK2

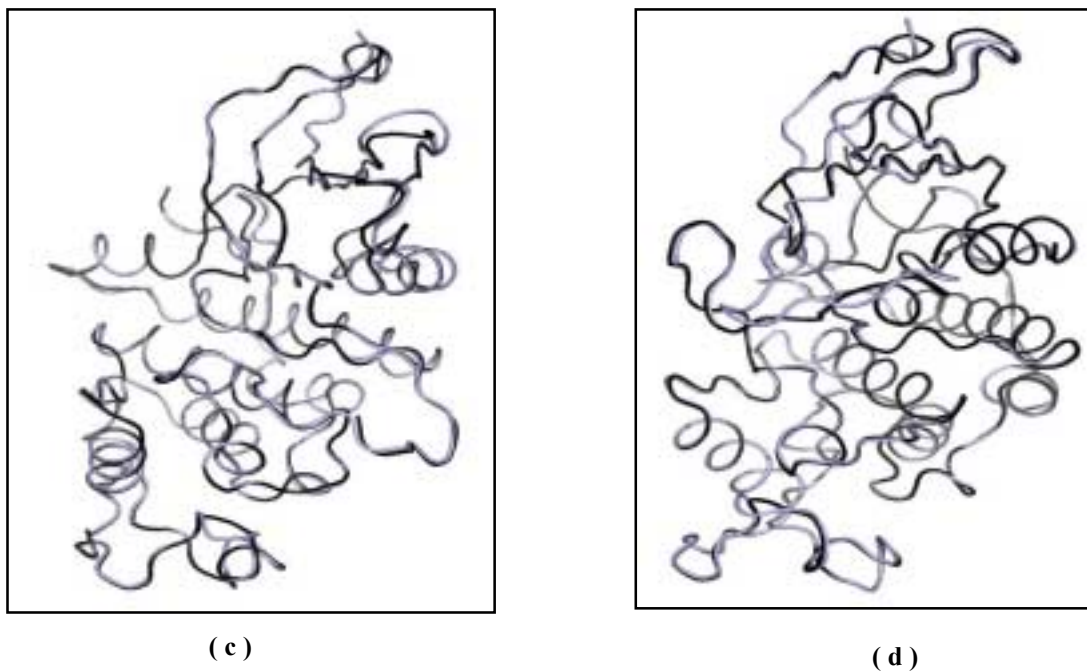
| MODELOS    | RMSD<br>sobreposição (Å) |
|------------|--------------------------|
| CDK10: ATP | 0,493                    |
| CDK10: flp | 0,080                    |
| CDK10: olo | 0,094                    |
| CDK10: rsc | 0,095                    |



( a )



( b )



**Figura 4-13** Sobreposições da CDK10 (linha azul) com seu molde CDK2 (linha preta), em ( a ) CDK10:ATP com CDK2:ATP, em ( b ) CDK10:flp com CDK2:flp, ( c ) CDK10:olo com CDK2:olo e em ( d ) CDK10:rsc com CDK2:rsc .

O LSQKAB gera também um arquivo com as coordenadas atômicas da sobreposição. Este arquivo de coordenadas atômicas foi utilizado como entrada no programa XtalView (McREE, 1999) para visualizar e analisar estas estruturas (modelos gerados) com base em sua sobreposição.

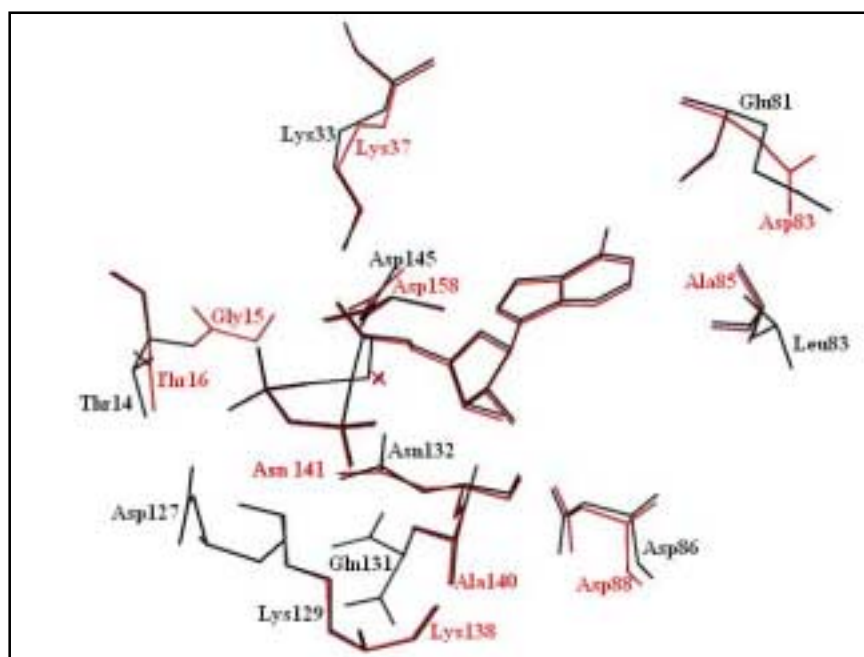
#### 4.4. Interação CDK- Inibidores

Os níveis de especificidade e afinidade entre uma enzima e seu inibidor, dependem de interações iônicas, ligações de hidrogênio direcionais e da complementariedade das superfícies de contato (CANDURI et al., 2001; DE AZEVEDO et al., 2002; KIM et al., 1996).

A seguir as (Tabelas 4-11 a 4-14), relacionam as interações por ligações de hidrogênio entre a CDK2 e CDK8 e seus 3 inibidores, além do ATP (seu ligante natural), juntamente com as (Figuras 4-14 a 4-17) que mostram os sítios de interação dos ligantes com as CDK2 e CDK8.

**Tabela 4-11** Ligações de Hidrogênio intermoleculares entre as proteínas CDK2 e CDK8 e o ATP.

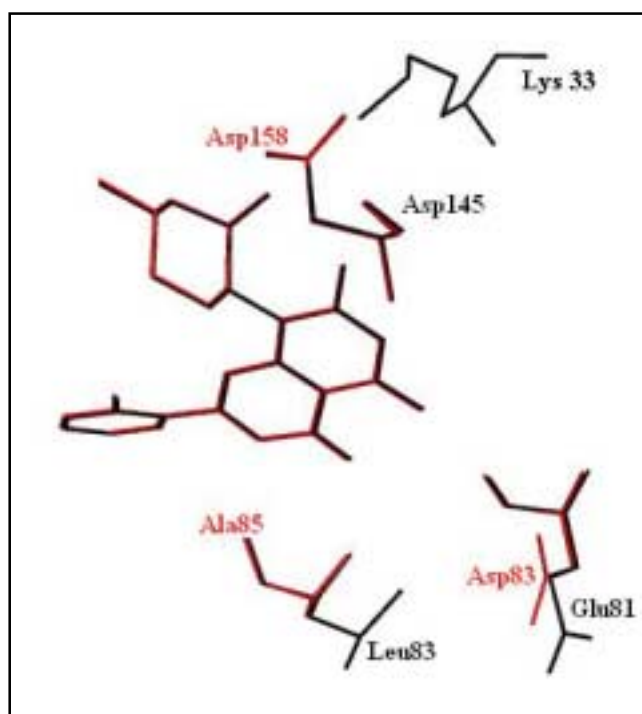
| CDK2 (molde) | Ligações de Hidrogênio (Å) | ATP | CDK8 (modelo) | Ligações de Hidrogênio (Å) | ATP |
|--------------|----------------------------|-----|---------------|----------------------------|-----|
| Thr 14 N     | 2,84                       | O2G | Gly 15 N      | 3,28                       | O2A |
| Thr 14 OG1   | 3,29                       | O2G | Thr 16 N      | 2,88                       | O2G |
| Thr 14 OG1   | 2,86                       | O3G | Thr 16 OG1    | 3,03                       | O3G |
| Lys 33 NZ    | 2,91                       | O1A | Lys 37 NZ     | 2,92                       | O1A |
| Glu 81 O     | 2,79                       | N6  | Lys 37 NZ     | 3,22                       | O2A |
| Leu 83 N     | 3,28                       | N1  | Asp 83 O      | 2,77                       | N6  |
| Asp 86 OD2   | 2,94                       | O3  | Ala 85 N      | 3,26                       | N1  |
| Asp 127 OD2  | 3,16                       | O3G | Asp 88 OD2    | 2,85                       | O3  |
| Lys 129 NZ   | 3,16                       | O3B | Asp 136 OD2   | 3,15                       | O1A |
| Lys 129 NZ   | 2,97                       | O3G | Lys 138 NZ    | 3,09                       | O3B |
| Gln 131 OE1B | 2,97                       | O1B | Lys 138 NZ    | 2,96                       | O3G |
| Gln 131 NE2B | 3,01                       | O1B | Ala 140 O     | 2,57                       | O2  |
| Gln 131 O    | 2,51                       | O2  | Asn 141 OD1   | 2,64                       | O2B |
| Asn 132 OD1  | 2,81                       | O2B | Asn 141 OD1   | 3,25                       | O1G |
| Asn 132 OD1  | 2,93                       | O1G | Asp 158 OD2   | 2,86                       | O1A |
| Asp 145 OD2  | 2,73                       | O1A | Asp 158 OD2   | 2,77                       | O1G |
| Asp 145 OD2  | 2,57                       | O1G |               |                            |     |



**Figura 4-14** Sobreposição dos sítios de ligação da CDK2 em preto e da CDK8 em vermelho, na presença de ATP.

**Tabela 4-12** Ligações de Hidrogênio intermoleculares entre as proteínas CDK2 e CDK8 e o Flavopiridol .

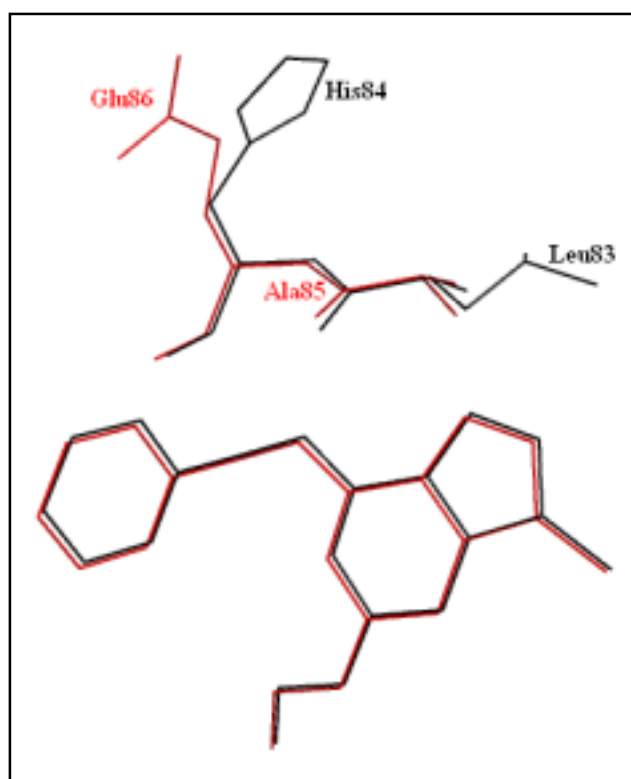
| CDK2 (molde) | Ligações de Hidrogênio (Å) | Flavopiridol | CDK8 (modelo) | Ligações de Hidrogênio (Å) | Flavopiridol |
|--------------|----------------------------|--------------|---------------|----------------------------|--------------|
| Lys 33 NZ    | 3,03                       | O3           | Asp 83 O      | 2,86                       | O5           |
| Glu 81 O     | 2,88                       | O5           | Ala 85        | 2,74                       | O4           |
| Leu 83       | 2,73                       | O4           | Asp 158 OD2   | 3,21                       | O3           |
| Asp 145 OD2  | 3,19                       | O3           |               |                            |              |



**Figura 4-15** Sobreposição dos sítios de ligação da CDK2 em preto e da CDK8 em vermelho, na presença de Flavopiridol .

**Tabela 4-13** Ligações de Hidrogênio intermoleculares entre as proteínas CDK2 e CDK8 e a Olomoucina

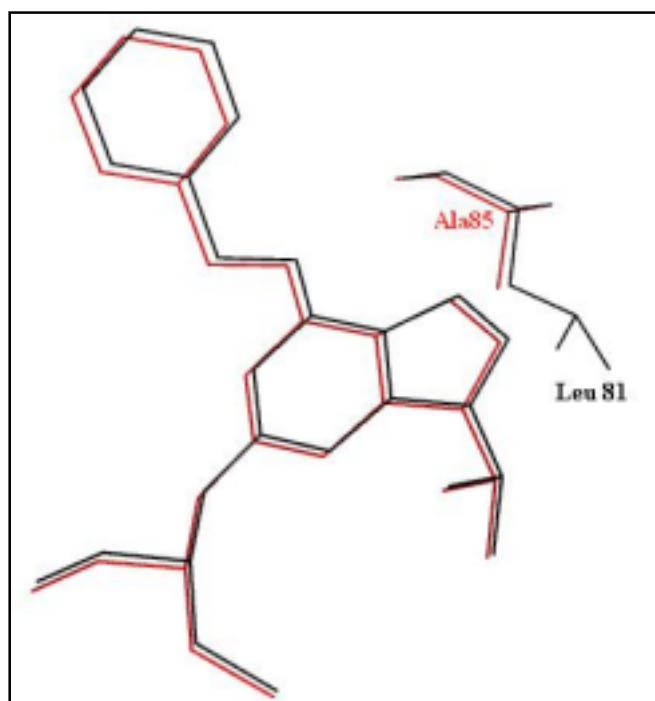
| CDK2 (molde) | Ligações de Hidrogênio (Å) | Olomoucina | CDK8 (modelo) | Ligações de Hidrogênio (Å) | Olomoucina |
|--------------|----------------------------|------------|---------------|----------------------------|------------|
| Leu 83 O     | 2,03                       | N6         | Ala 85 O      | 2,37                       | N6         |
| Leu 83 N     | 2,29                       | N7         | Ala 85 N      | 2,49                       | N7         |
| His 84 O     | 3,46                       | N6         | Glu 86 O      | 2,61                       | N6         |



**Figura 4-16** Sobreposição dos sítios de ligação da CDK2 em preto e da CDK8 em vermelho, na presença de Olomoucina .

**Tabela 4-14** Ligações de Hidrogênio intermoleculares entre as proteínas CDK2 e CDK8 e a Roscovitina.

| CDK2 (molde) | Ligações de Hidrogênio (Å) | Roscovitina | CDK8 (modelo) | Ligações de Hidrogênio (Å) | Roscovitina |
|--------------|----------------------------|-------------|---------------|----------------------------|-------------|
| Ala 85 O     | 2,86                       | N6          | Leu 81 O      | 2,82                       | N6          |
| Ala 85 N     | 3,38                       | N7          | Leu 81 N      | 3,38                       | N7          |

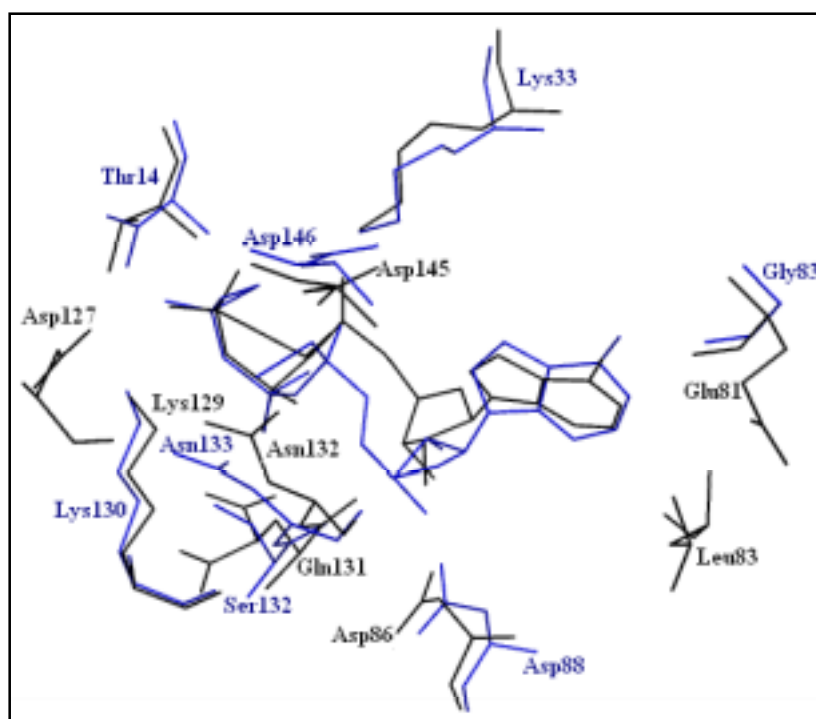


**Figura 4-17** Sobreposição dos sítios de ligação da CDK2 em preto e da CDK8 em vermelho, na presença de Roscovitina.

As (Tabelas 4-15 a 4-18), relacionam as interações por ligações de hidrogênio entre a CDK2 e CDK10 e seus 3 inibidores, além do ATP (seu ligante natural), juntamente com as (Figuras 4-18 a 4-21) que mostram a interação dos resíduos do sítio e os ligantes, com as CDK2 e CDK10.

**Tabela 4-15** Ligações de Hidrogênio intermoleculares entre as proteínas CDK2 e CDK10 e o ATP.

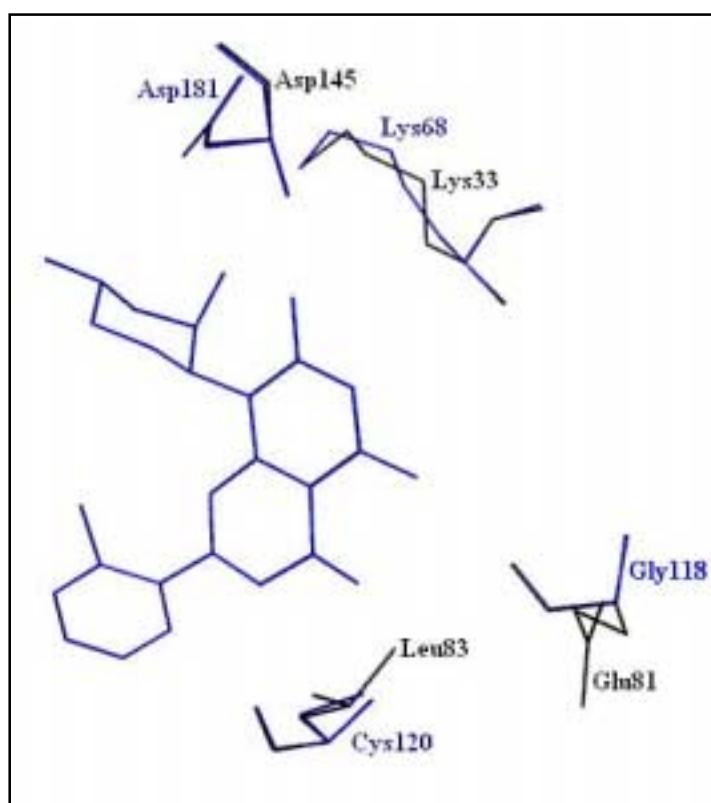
| CDK2 (molde) | Ligações de Hidrogênio (Å) | ATP | CDK10 (modelo) | Ligações de Hidrogênio (Å) | ATP |
|--------------|----------------------------|-----|----------------|----------------------------|-----|
| Thr 14 N     | 2,84                       | O2G | Thr 14 N       | 3,02                       | O1G |
| Thr 14 OG1   | 3,29                       | O2G | Thr 14 N       | 3,14                       | O3G |
| Thr 14 OG1   | 2,86                       | O3G | Thr 14 OG1     | 2,65                       | O3G |
| Lys 33 NZ    | 2,91                       | O1A | Lys 33 NZ      | 2,78                       | O1A |
| Glu 81 O     | 2,79                       | N6  | Gly 83 O       | 2,79                       | N6  |
| Leu 83 N     | 3,28                       | N1  | Asp 88 OD2     | 2,77                       | O3  |
| Asp 86 OD2   | 2,94                       | O3  | Lys 130 NZ     | 3,12                       | O3B |
| Asp 127 OD2  | 3,16                       | O3G | Lys 130 NZ     | 3,19                       | O3G |
| Asp 129 NZ   | 3,16                       | O3B | Ser 132 OG     | 2,78                       | O1B |
| Lys 129 NZ   | 2,97                       | O3G | Asn 133 OD1    | 3,01                       | O3A |
| Gln 131 OE1B | 2,97                       | O1B | Asp 146 OD2    | 3,11                       | O1A |
| Gln 131 NE2B | 3,01                       | O1B | Asp 146 OD2    | 2,45                       | O2A |
| Gln 131 O    | 2,51                       | O2  | Asp 146 OD2    | 3,26                       | O2G |
| Asn 132 OD1  | 2,81                       | O2B |                |                            |     |
| Asn 132 OD1  | 2,93                       | O1G |                |                            |     |
| Asp 145 OD2  | 2,73                       | O1A |                |                            |     |
| Asp 145 OD2  | 2,57                       | O1G |                |                            |     |



**Figura 4-18** Sobreposição dos sítios de ligação da CDK2 em preto e da CDK10 em azul, na presença de ATP.

**Tabela 4-16** Ligações de Hidrogênio intermoleculares entre as proteínas CDK2 e CDK10 e o Flavopiridol.

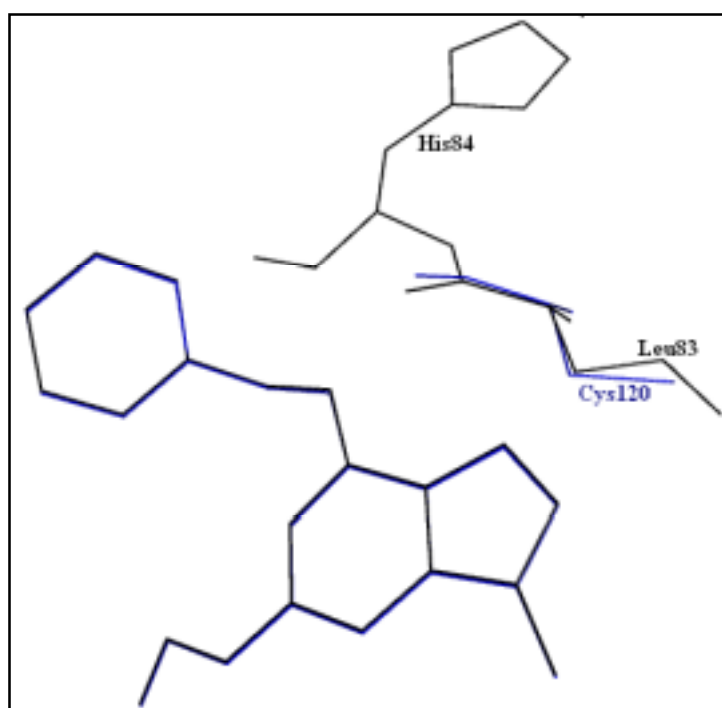
| CDK2 (molde) | Ligações de Hidrogênio (Å) | Flavopiridol | CDK10 (modelo) | Ligações de Hidrogênio (Å) | Flavopiridol |
|--------------|----------------------------|--------------|----------------|----------------------------|--------------|
| Lys 33 NZ    | 3,03                       | O3           | Lys 68 NZ      | 3,02                       | O3           |
| Glu 81 O     | 2,88                       | O5           | Gly 118 O2     | 2,87                       | O5           |
| Leu 83       | 2,73                       | O4           | Cys 120 N      | 2,74                       | O4           |
| Asp 145 OD2  | 3,19                       | O3           | Asp 181 OD2    | 3,20                       | O3           |



**Figura 4-19** Sobreposição dos sítios de ligação da CDK2 em preto e da CDK10 em azul, na presença de Flavopiridol.

**Tabela 4-17** Ligações de Hidrogênio intermoleculares entre as proteínas CDK2 e CDK10 e a Olomoucina.

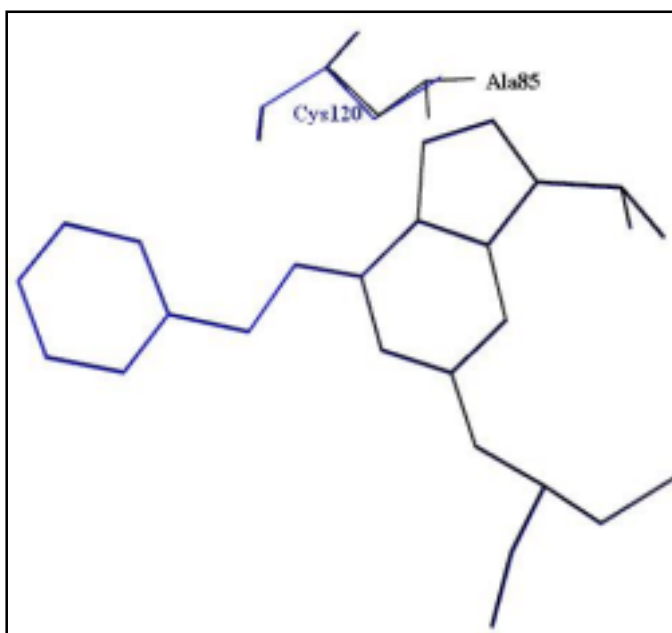
| CDK2 (molde) | Ligações de Hidrogênio (Å) | Olomoucina | CDK10 (modelo) | Ligações de Hidrogênio (Å) | Olomoucina |
|--------------|----------------------------|------------|----------------|----------------------------|------------|
| Leu 83 O     | 2,03                       | N6         | Cys 120 O      | 2,36                       | N6         |
| Leu 83 N     | 2,29                       | N7         | Cys 120 N      | 2,48                       | N7         |
| His 84 O     | 3,46                       | N6         |                |                            |            |



**Figura 4-20** Sobreposição dos sítios de ligação da CDK2 em preto e da CDK10 em azul, na presença de Olomoucina.

**Tabela 4-18** Ligações de Hidrogênio intermoleculares entre as proteínas CDK2 e CDK10 e a roscovitina.

| CDK2 (molde) | Ligações de Hidrogênio (Å) | Roscovitina | CDK10 (modelo) | Ligações de Hidrogênio (Å) | Roscovitina |
|--------------|----------------------------|-------------|----------------|----------------------------|-------------|
| Ala 85 O     | 2,86                       | N6          | Cys 120 O      | 2,83                       | N6          |
| Ala 85 N     | 3,38                       | N7          | Cys 120 N      | 3,38                       | N7          |



**Figura 4-21** Sobreposição dos sítios de ligação da CDK2 em preto e da CDK10 em azul, na presença de Roscovitina.

Os valores da área de contato são apresentados na (Tabela 4-19):

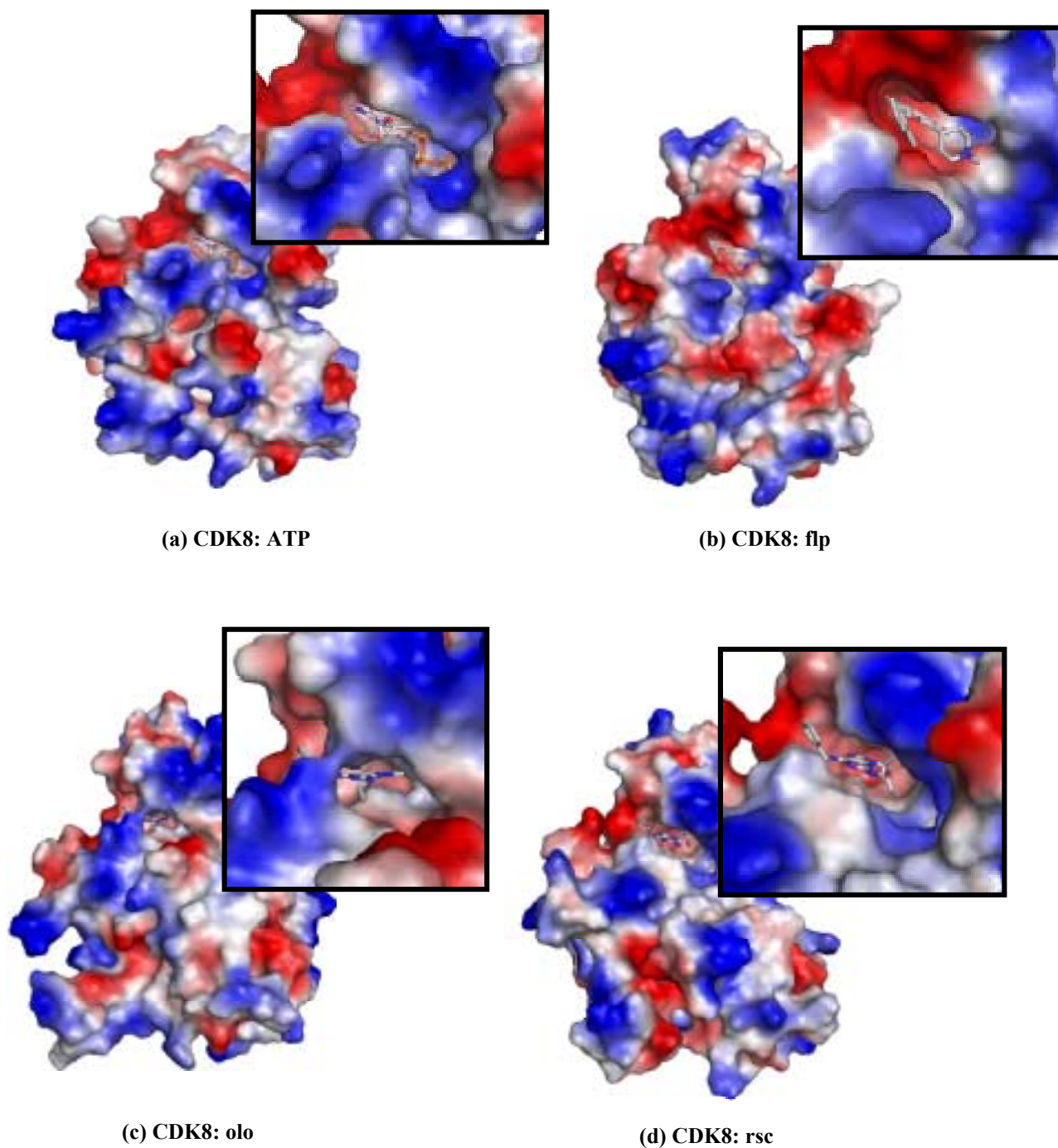
**4-19** Tabela com os valores da área de contato para cada tipo de inibidor e sua respectiva proteína.

| Inibidor     | CDK2 | CDK8 | CDK10 |
|--------------|------|------|-------|
| ATP          | 316  | 308  | 300   |
| Flavopiridol | 320  | 305  | 322   |
| Olomoucina   | 288  | 270  | 275   |
| Roscovitina  | 357  | 323  | 412   |

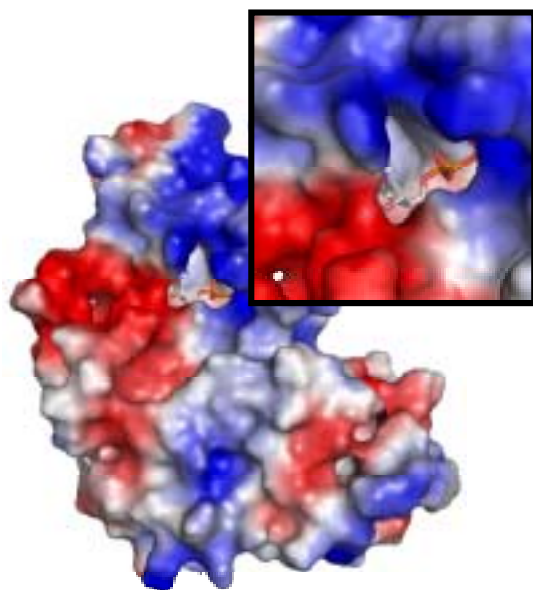
- Os valores da área de contato são dados em Å<sup>2</sup>.

Estes resultados foram obtidos com a execução das rotinas do pacote CCP4: areaimol, diffarea e resarea (CCP4, 1994).

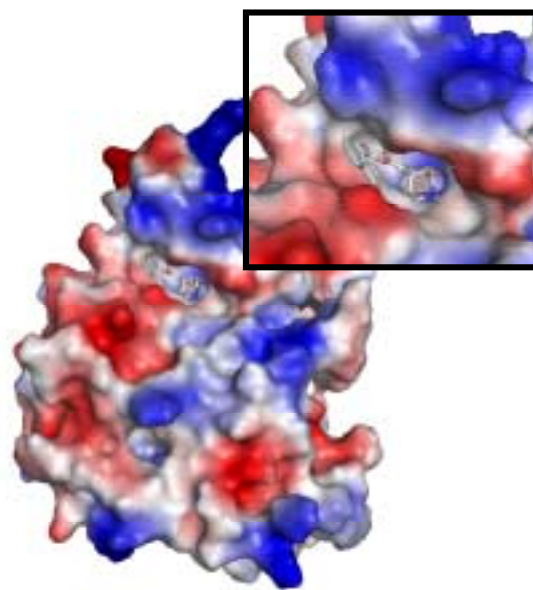
A análise do potencial eletrostático de superfície revelou a presença de certa complementaridade de cargas na região de contato entre as CDKs e seus inibidores, geradas pelo programa PYMOL (DELANO, 2004) (Figura 4-22) para a CDK8, e (Figura 4-23) para a CDK10.



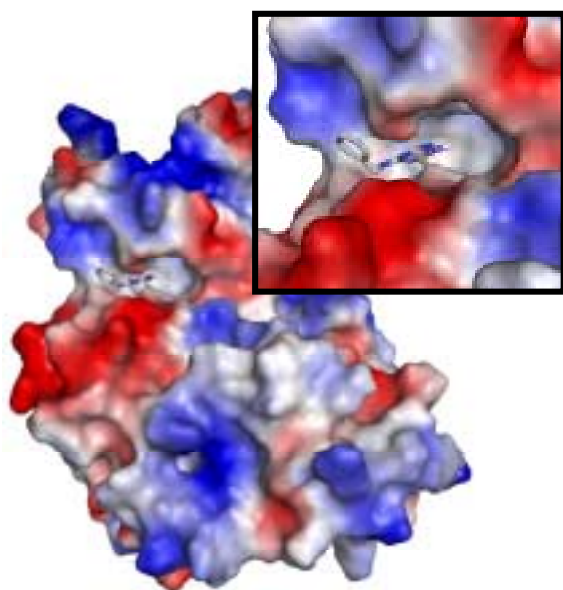
**Figura 4-22-** Potenciais Eletrostáticos dos quatro complexos da CDK8



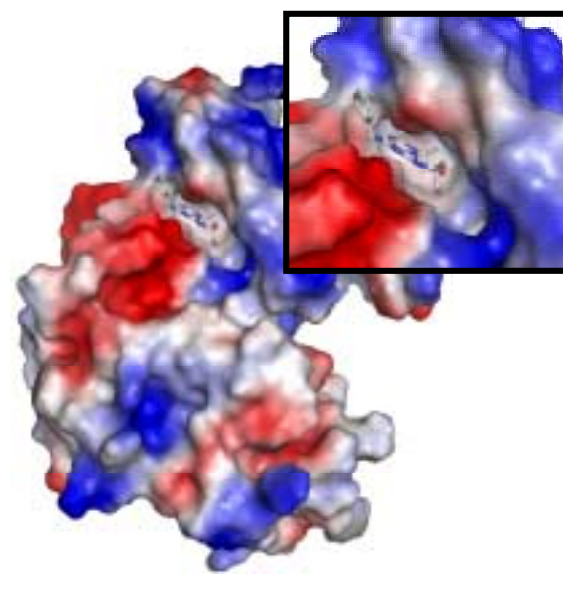
(a) CDK10: ATP



(a) CDK10: flp



(a) CDK10: olo



(a) CDK10: rsc

**Figura 4-23-** Potenciais Eletrostáticos dos quatro complexos da CDK10

## 5. CONCLUSÃO

Em relação às estruturas das CDKs, as modelagens obtidas dos complexos CDK8 e CDK10 e seus ligantes, indicam que estes apresentam semelhanças com seu molde CDK2; isso já era esperado, pois seus alinhamentos mostraram identidades superiores a 30%. Percebe-se pequenas variações quanto aos domínios das estruturas bilobais, quando comparadas a CDK2. Já os sítios de ligações se conservam exatamente na mesma posição do molde, ou seja, entre os domínios N e C terminais. Em relação aos ligantes, observando as tabelas de ligações de hidrogênio, percebe-se uma similaridade entre o número de ligações do molde e dos modelos, o que se conclui que os modelos em estudo possuem o mesmo comportamento do molde, quando ligados ao ATP, ou aos inibidores flavopiridol, olomoucina e roscovitina. Pode-se dizer que as pequenas variações estruturais e a diferença no número de ligações de hidrogênio dos modelos/ligantes e seu molde/ligantes são causadas pelas mudanças conformacionais de cadeias laterais que influenciam na formação de ligações e da área de contato, e estão relacionadas com o tamanho das seqüências da CDK8 e CDK10. Conclui-se, portanto, que pelos dados analisados, as modelagens das CDK8 e CDK10, são boas referências para o desenho de drogas que inibam a função das mesmas, principalmente, quando ligadas a roscovitina, que apresentou os melhores resultados para ambas CDKs. Dessa forma, o objetivo proposto para este trabalho foi alcançado com êxito.

## 6. REFERÊNCIAS BIBLIOGRÁFICAS

- ALBERTS, B. et al. **Biologia Molecular da Célula**. 3. ed. Porto Alegre: ArtMed, 1999,1294p.
- ALLEN, F. H. et al. The Cambridge Crystallographic Data Centre: computer-based search, retrieval, analysis and display of information. **Acta Crystallographica Section B**, v. 35, 1979, p.2331-2339.
- AKOULITCHEV, S.; CHUIKOV, S.; REINBERG, D. TFIID is negatively regulated by cdk8-containing mediator complexes. **Nature**, v. 407, 2000, p.102-106.
- ARCURI, H. A. et al; Molecular Models for Shikimate Pathway Enzymes of *Xylella fastidiosa*. **Biochemical Biophysical Research Communication**, v.320, n.3, 2004, p.979-991.
- BARVIAN, M. et al. Pyrido[2,3-d]pyrimidin-7-one inhibitors of cyclin-dependent kinases. **Journal of Medicinal Chemistry**, v. 43, 2000, p.4606-4616.
- BAYAT, A. Science, medicine and the future: Bioinformatics. **Journal Molecular Biology**, v.324, 2002, p. 1018-1022.
- BERMAN, H. M. et al. The Protein Data Bank. **Nucleic Acids Research**, v. 28, 2000, p. 235-242.
- BLUNDELL, T. L. et al. Knowledge-based prediction of protein structures and the design of novel molecules. **Nature**, v.326, n.6111, 1987, p. 347 – 352.
- BOWIE, J. U., et al; A method to identify protein sequences that fold into a known three-dimensional structure. **Science**, v.253, 1991, p.164-170.
- BROWNE, W.J. et al. A possible three-dimensional structure of bovine lactalbumin based on that of hen's egg-white lysozyme. **Journal Molecular Biology**, v.42, 1969, p.65-86.
- BRUNGER, A. T. **X-PLOR a system for X-ray Crystallography and NMR**. New Haven USA: Yale University Press, 1992.
- CANDURI, F. et al. Crystal structure of human uropepsin at 2,45 Å resolution. **Acta Crystallographica Section D**, v. 57, 2001, p. 1560-70.
- CANDURI, F.; DE AZEVEDO, W. F. JR. Structural basis for interaction of inhibitors with Cyclin-Dependent Kinase 2. **Current Computer-Aided Drug Design**, v.1, 2005, p.53-64.
- CLAESSENS, M. et al. Modelling the polypeptide backbone with 'spare parts' from known protein structures. **Protein Engineering**, v. 2, 1989, p. 335-345.
- Collaborative Computational Project N° 4., The CCP4 suite: programs for proteins crystallography. **Acta Crystallographica Section D**, v. 50, 1994, p.760-763.

CORPET, F. Multiple sequence alignment with hierarchical clustering. **Nucleic Acids Research**, v. 16, 1998, p.10881-10890.

DA SILVEIRA, N. J., et al. Structural bioinformatics study of PNP from *Schistosoma mansoni*. **Biochemical Biophysical Research Communications**, v.322, 2004, p.100-104.

DAÍ Y.; GRANT, S. Cyclin-dependent kinase inhibitors. **Current Opinion in Pharmacology**, v.3, 2003, p.362-370.

DE AZEVEDO, W.F. Jr. et al. Structural basis for specificity and potency of a flavonoid inhibitor of human CDK2, a cell cycle kinase. **Proceedings of the National Academy of Sciences of United State of America**, v. 93, n. 7, 1996, p. 2735-2740.

DE AZEVEDO JR., W. F.; CANDURI, F.; SILVEIRA, N. J. F. Structural basis for inhibition of cyclin-dependent kinase 9 by flavopiridol. **Biochemical and Biophysical Research Communications**. 293(1), 566-571, 2002.

DELANO, W. L. **The PyMOL Molecular Graphics System**. San Carlos: CA DeLano Scientific, 2004.

EDAMATSU, H. et al. Cdk inhibitors, roscovitine and olomoucine, synergize with farnesyltransferase inhibitor (FTI) to induce efficient apoptosis of human cancer cell lines. **Oncogene**, 19, 3059-3068, 2000.

ENGH, R. A.; HUBER, R. Accurate bond and angle parameters for X-ray protein structure refinement. **Acta Crystallographica**, v.A47, 1991, p.392-400.

ESWAR, N. et al. Tools for comparative protein structure modeling and analysis. **Nucleic Acids Research**. v.31, 2003, p.3375-380.

FILHO, O. A. S.; ALENCASTRO, R. B. Modelagem de Proteínas por Homologia. **Química Nova**, v.26, n.2, 2003, p.253-259 2003.

FISER, A.; DO, R. K.; SALI, A. Modelling of loops in protein structures, **Protein Science**, v. 9, 2000, p. 1753-1773.

FISER, A. et al; Evolution and Physics in Comparative Protein Structure Modeling. **Accounts of Chemical Research**, v. 35, 2002, p.413-421.

FURET, P. et al. Structure-based design of potent CDK1 inhibitors derived from olomoucine. **Journal of Computer-Aided Molecular Design**, v. 14, p. 403-409, 2000.

GIBAS, C.; JAMBECK, P. **Desenvolvendo Bioinformática**, Rio de Janeiro: Campus, 2001, p.179-180.

GREER, J. Comparative model-building of the mammalian serine proteases. **Journal Molecular Biology**, v.153, 1981, p.1027-1042.

GREER, J. Comparative modeling methods: application to the family of the mammalian serine proteases. **Proteins**, v. 7, n. 4, 1990, p. 317-334.

GOULD, K. L. Cyclin-dependent protein kinases. In Woodgett, J. **Protein Kinase Functions**. Oxford: Hames, B. D.; Glover, D. M., 2000, p. 277-302.

HOOFT, R. W. et al. Errors in protein structures. **Nature**, v.381, 1996, p.272.

JOHNSON, M. S. et al. Knowledge-based protein modelling. **Critical Reviews in Biochemistry and Molecular Biology**, v.219, 1994, p.1-68.

JUNQUEIRA, L.C.; CARNEIRO, J. **Biologia Celular e Molecular**. 6. ed. São Paulo: Guanabara Koogan, 1997, 300p.

KABASCH, W.; SANDER, C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. **Biopolymers**, v. 22, n.12, 1983, p. 2577-2637.

KEARSLEY, S. K. On the orthogonal transformation used for structural comparisons. **Acta Crystallographica**, v. A45, 1989, p.208-210.

KIM, S. H.; SCHULZE-GAHMEN, U.; BRANDSEN, J. Structural basis for chemical inhibitor of CDK2. **Progress in Cell Cycle Research**, v. 2, 1996, p. 137-145.

KNOCKAERT, M.; GREENGARD, P.; MEIJER, L. Pharmacological inhibitors of cyclin-dependent kinases. **Trends in Pharmacological Sciences**, v. 23, 2002, p. 417-425.

LASKOWSKI, R. A. et al. PROCHECK - A program to check the stereochemical quality of protein structures. **Journal. Applied Crystallography**, v. 26, 1993, p. 283-291.

LEVITT, M. Accurate modeling of protein conformation by automatic segment matching. **Journal Molecular Biology**, v. 226, 1992, p. 507-533.

LODISH, H. et al. **Molecular Cell Biology**. 3. ed. New York: Scientific American Books, 1995, 1400p.

LUTHY, R. et al. Assessment of protein models with three-dimensional profiles. **Nature**, v. 356, 1992, p. 383-385.

KABSCH, W.; SANDER, C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. **Biopolymers**, v. 22, n. 12, 1983, p. 2577-2637.

MANHANI, K. K. et al. Molecular models of protein kinase 6 from *Plasmodium falciparum*. **Journal Molecular Model**, v. 12, 2005, p. 42-8.

MARTI-RENO, M. A. et al; Comparative protein structure modeling of genes and genomes. **Annual Review of Biophysics and Biomolecular Structure**, v. 29, 2000, p. 291-325.

MARTI-RENO, M. A.; MADHUSUDHAN, M. S.; SALI, A. Alignment of protein sequences by their profiles, **Protein Science**, v. 13, 2004, p. 1071-1087.

McREE, D. E. XtalView/Xfit – A versatile Program for Manipulating Atomic Coordinates and Electron Density. **Journal Structural Biology**, v.125, 1999, p.156-165.

MEIJER, L. Cyclin-dependent kinases inhibitors as potential anticancer, antineurodegenerative, antiviral and antiparasitic agents. **Drug Resistance Updates**, v. 3, n. 2, 2000, p. 83-88.

MEIJER, L.; RAYMOND, E. Roscovitine and other purines as kinase inhibitors. From starfish Oocytes to clinical trials. **Accounts of Chemical Research**, v. 36, 2003, p. 417-425.

MORRIS, A. L et al. Stereochemical quality of protein structure coordinates. **Journal Applied Crystallography**, v. 12, 1986, p. 345-364.

MORRIS, A. L. et al. Stereochemical quality of protein structure coordinates. **Proteins**, v.12, 1992, p.345-364.

PEREIRA, J. H. et al. Structural bioinformatics study of EPSP synthase from *Mycobacterium tuberculosis*. **Biochemical Biophysical Research Communications**, v.312, 2003, p.608-614.

RAMACHANDRAN, G. N.; RAMAKRISHNAN, C.; SASISSEKHARAN, V. Stereochemistry of polypeptide chain configurations. **Journal Molecular Biology**, v. 7, 1963, p. 95-99.

RAMAKRISHNAN, C.; RAMACHANDRAN, G. N. Stereochemical criteria for polypeptide and protein chain conformations. II. Allowed conformations for a pair of peptide units. **Biophysical Journal**, v. 5, n. 6, 1965, p. 909-933.

SALI, A; BLUNDELL, T. L. Comparative protein modelling by satisfaction of spatial restraints. **Journal Molecular Biology**, v.234, 1993, p.779-815.

SALI, A. et al. Evaluation of comparative protein modeling by MODELLER. **Proteins** v.23, 1995, p.318-326.

SANCHEZ, R.; SALI, A. Advances in comparative protein-structure modelling. **Current Opinion in Structural Biology**, v.7, 1997, p.206-214.

SANCHEZ, R.; SALI, A. Large-scale protein structure modeling of the *Saccharomyces cerevisiae* genome. **Proceedings of the National Academy of Sciences of United State of America**, v. 95, 1998, p. 13597-13602.

SANCHEZ, R.; SALI, A; Comparative protein structure modeling. Introduction and practical examples with modeler. **Methods in Molecular Biology**, v. 143, 2000, p. 97-129.

SAUSVILLE, E.A. Complexities in the development of cyclin-dependent-kinase inhibitor drugs. **Trends in Molecular Medicine**, v. 8., 2002, p. S32-S37.

SCHWIETERS, C.D. et al. The Xplor-NIH NMR Molecular Structure Determination Package. **Journal of Magnetic. Resonance.**, v. 160, 2003, p. 66-74.

SENDEROWICZ, A.M.; SAUSVILLE, E.A. Preclinical and clinical development of cyclin-dependent kinase modulators. **Journal of the National Cancer Institute**, v. 92, 2000, p. 376-387.

SIELECHI, T. M. et al. Cyclin-dependent kinase inhibitors: Useful targets in cell cycle regulation. **Journal of Medicinal Chemistry**, v. 43, n. 1, 2000, p. 1-16.

SONI, R. et al. Selective in vivo and in vitro effects of a small molecule inhibitor of cyclin-dependent kinase 4. **Journal of the National Cancer Institute**, v. 21, 2001, p. 436-446.

SCHUTTE, B. et al. The effect of the cyclin-dependent kinase inhibitor olomoucine on cell cycle kinetics. **Experimental Cell Research**, v. 236, 1997, p.4-15.

TRAMONTANO, A. et al. Analysis and assessment of comparative modeling predictions in CASP4. **Proteins**, v. 45, 2001, p. 22-38.

UCHOA, H. B. et al. PARMODEL: a web server for automated comparative modeling of proteins. **Biochemical Biophysical Research Communications**, v. 325, n. 4, 2004, p. 1481-1486.

VAN GUNSTEREN, W. F. et al. Biomolecular Simulations: The GROMOS96 Manual and User Guide. **VdF Hochschulverlag ETHZ**, Zurich, 1996.

VRIEND, G. WHAT IF: A molecular modeling and drug design program. **Journal of Molecular Graphics**, v.8, 1990, p.52-56.

WESTBROOK, J. et al. The Protein Data Bank: unifying the archive. **Nucleic Acids Research**, v. 30, 2002, p. 245-248.