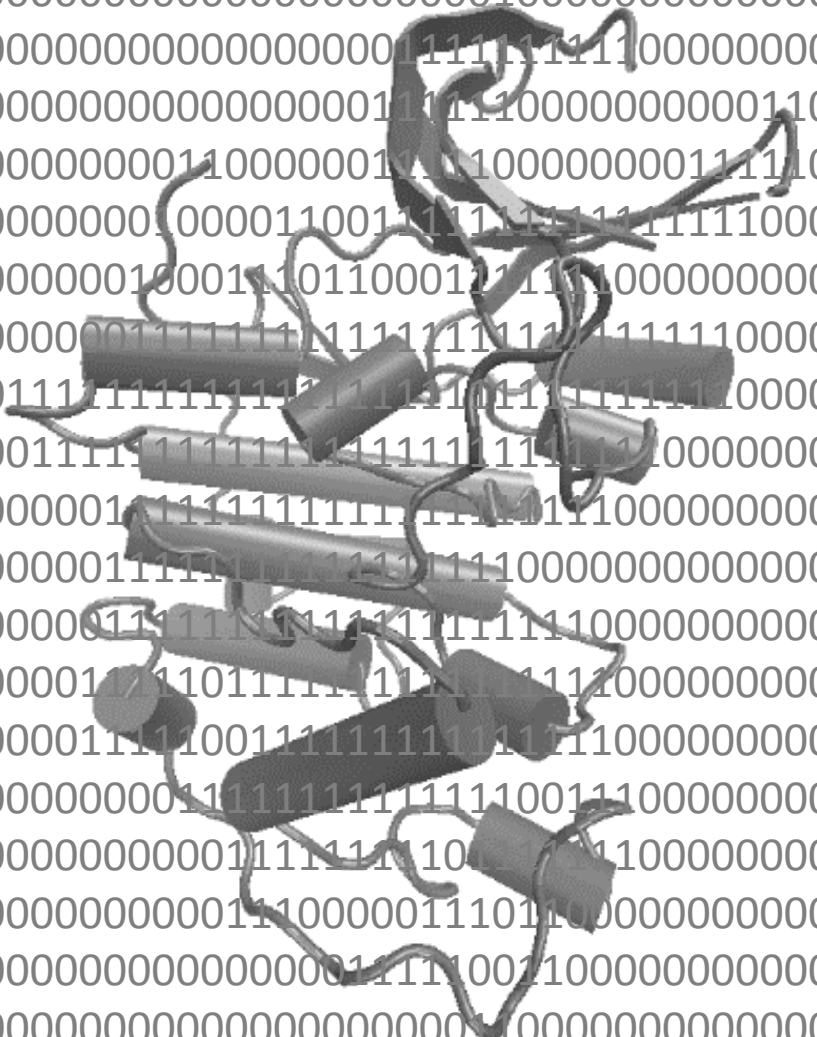


Biologia Estrutural

Qualidade de Modelos Estruturais de Proteínas



Como toda informação experimental, os dados de cristalografia apresentam limitações intrínsecas da técnica, notadamente a resolução cristalográfica. Tais limitações resultam em imprecisões nas posições atômicas da estrutura final. Os erros nas posições atômicas podem ser dimensionados a partir da comparação com valores ideais das distâncias interatômicas, ângulos de ligação e ângulos de torção. Abaixo temos a parte inicial do arquivo PDB para a estrutura da chiquimato quinase de *Mycobacterium tuberculosis* (código de acesso PDB: 1WE2) (Pereira *et al*, 2004). As coordenadas atômicas estão em destaque. Essas informações que são usadas para avaliar a qualidade de informação estrutural depositada.

ATOM	1	N	ALA	A	2	26.424	27.695	8.672	1.00	38.67	N
ATOM	2	CA	ALA	A	2	26.646	26.892	9.925	1.00	38.14	C
ATOM	3	C	ALA	A	2	28.080	27.086	10.395	1.00	36.44	C
ATOM	4	O	ALA	A	2	29.031	26.973	9.615	1.00	36.87	O
ATOM	5	CB	ALA	A	2	26.357	25.374	9.691	1.00	37.83	C
ATOM	6	N	PRO	A	3	28.256	27.322	11.685	1.00	35.89	N
ATOM	7	CA	PRO	A	3	29.592	27.533	12.232	1.00	32.15	C
ATOM	8	C	PRO	A	3	30.578	26.368	12.108	1.00	31.97	C
ATOM	9	O	PRO	A	3	30.179	25.186	12.111	1.00	32.12	O
ATOM	10	CB	PRO	A	3	29.292	27.869	13.700	1.00	31.84	C
ATOM	11	CG	PRO	A	3	27.920	28.471	13.656	1.00	31.93	C
ATOM	12	CD	PRO	A	3	27.228	27.520	12.720	1.00	33.67	C

Pereira JH, de Oliveira JS, Canduri F, Dias MV, Palma MS, Basso LA, Santos DS, de Azevedo WF Jr. Structure of shikimate kinase from *Mycobacterium tuberculosis* reveals the binding of shikimic acid. *Acta Crystallogr D Biol Crystallogr*. 2004; 60(Pt 12 Pt 2):2310-9.

Ao depositamos as coordenadas atômicas de uma estrutura de macromolécula biológica junto ao *Protein Data Bank*, realizamos uma validação desta. Neste processo são calculados desvios entre a estrutura a ser depositada e os valores considerados ideais. Tais valores são expressos como o desvio médio quadrático (*root-mean square deviation, RMSD*). Veremos no presente texto, como os parâmetros geométricos são calculados, a partir das coordenadas atômicas da estrutura.

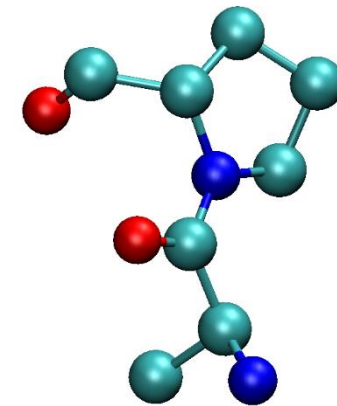
Tabela com valores de RMSD para a estrutura da chiquimato quinase de *Mycobacterium tuberculosis* (código de acesso no PDB: 1WE2).

Temperature Factor Modeling	
Mean Isotropic B Value	35.1
RMS Deviations	
Parameter Type	Deviation from Ideal
x_improper_angle_d	1.766
x_dihedral_angle_d	22.125
x_angle_deg	1.905
x_bond_d	0.017
Coordinate Error	
Luzzati ESD(Observed)	0.28
Luzzati Resolution Cutoff(Low)	6.0
Number of Non-Hydrogen Atoms Used in Refinement	
Protein Atoms	1226
Nucleic Acid Atoms	0
Heterogen Atoms	42
Solvent Atoms	144

A figura mostrada abaixo traz a estrutura tridimensional dos resíduos de aminoácido alanina 2 e prolina 3, da chiquimato quinase (código de acesso PDB: 1WE2). O arquivo PDB traz as coordenadas atômicas que foram usadas para gerar a figura. Para cada átomo mostrado na figura, temos uma posição atômica no arquivo PDB. Lembre-se, as coordenadas atômicas dos arquivos PDB estão em angstrom (Å), $1 \text{ \AA} = 10^{-10} \text{ m}$.

Trecho do arquivo PDB para os resíduos Ala2 e Pro3

ATOM	1	N	ALA	A	2	26.424	27.695	8.672	1.00	38.67	N
ATOM	2	CA	ALA	A	2	26.646	26.892	9.925	1.00	38.14	C
ATOM	3	C	ALA	A	2	28.080	27.086	10.395	1.00	36.44	C
ATOM	4	O	ALA	A	2	29.031	26.973	9.615	1.00	36.87	O
ATOM	5	CB	ALA	A	2	26.357	25.374	9.691	1.00	37.83	C
ATOM	6	N	PRO	A	3	28.256	27.322	11.685	1.00	35.89	N
ATOM	7	CA	PRO	A	3	29.592	27.533	12.232	1.00	32.15	C
ATOM	8	C	PRO	A	3	30.578	26.368	12.108	1.00	31.97	C
ATOM	9	O	PRO	A	3	30.179	25.186	12.111	1.00	32.12	O
ATOM	10	CB	PRO	A	3	29.292	27.869	13.700	1.00	31.84	C
ATOM	11	CG	PRO	A	3	27.920	28.471	13.656	1.00	31.93	C
ATOM	12	CD	PRO	A	3	27.228	27.520	12.720	1.00	33.67	C



Abaixo temos o detalhamento da estrutura, onde vemos as coordenadas atômicas usadas para gerar as posições para cada átomo da figura, exceto os hidrogênios. As coordenadas atômicas estão em angstrom (Å). A legenda de cores dos átomos está mostrada abaixo à direita.

ATOM	12	CD	PRO A	3	27.228	27.520	12.720	1.00	33.67	C
------	----	----	-------	---	--------	--------	--------	------	-------	---

ATOM	11	CG	PRO A	3	27.920	28.471	13.656	1.00	31.93	C
------	----	----	-------	---	--------	--------	--------	------	-------	---

ATOM	10	CB	PRO A	3	29.292	27.869	13.700	1.00	31.84	C
------	----	----	-------	---	--------	--------	--------	------	-------	---

ATOM	8	C	PRO A	3	30.578	26.368	12.108	1.00	31.97	C
------	---	---	-------	---	--------	--------	--------	------	-------	---

ATOM	9	O	PRO A	3	30.179	25.186	12.111	1.00	32.12	O
------	---	---	-------	---	--------	--------	--------	------	-------	---

ATOM	7	CA	PRO A	3	29.592	27.533	12.232	1.00	32.15	C
------	---	----	-------	---	--------	--------	--------	------	-------	---

ATOM	6	N	PRO A	3	28.256	27.322	11.685	1.00	35.89	N
------	---	---	-------	---	--------	--------	--------	------	-------	---

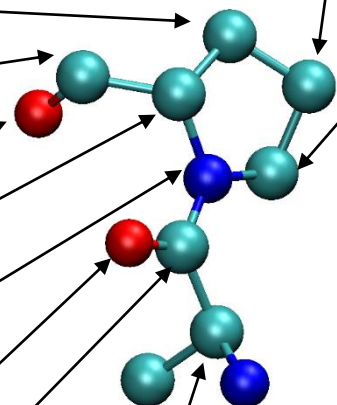
ATOM	4	O	ALA A	2	29.031	26.973	9.615	1.00	36.87	O
------	---	---	-------	---	--------	--------	-------	------	-------	---

ATOM	3	C	ALA A	2	28.080	27.086	10.395	1.00	36.44	C
------	---	---	-------	---	--------	--------	--------	------	-------	---

ATOM	5	CB	ALA A	2	26.357	25.374	9.691	1.00	37.83	C
------	---	----	-------	---	--------	--------	-------	------	-------	---

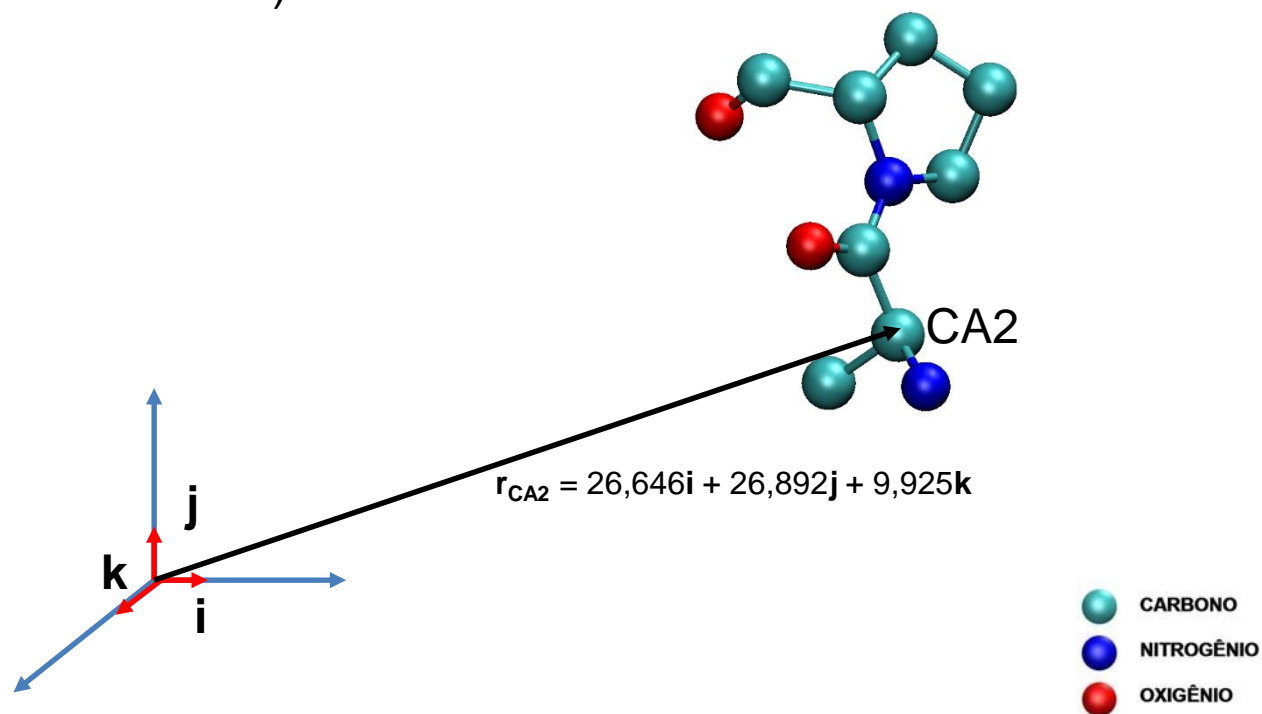
ATOM	2	CA	ALA A	2	26.646	26.892	9.925	1.00	38.14	C
------	---	----	-------	---	--------	--------	-------	------	-------	---

ATOM	1	N	ALA A	2	26.424	27.695	8.672	1.00	38.67	N
------	---	---	-------	---	--------	--------	-------	------	-------	---

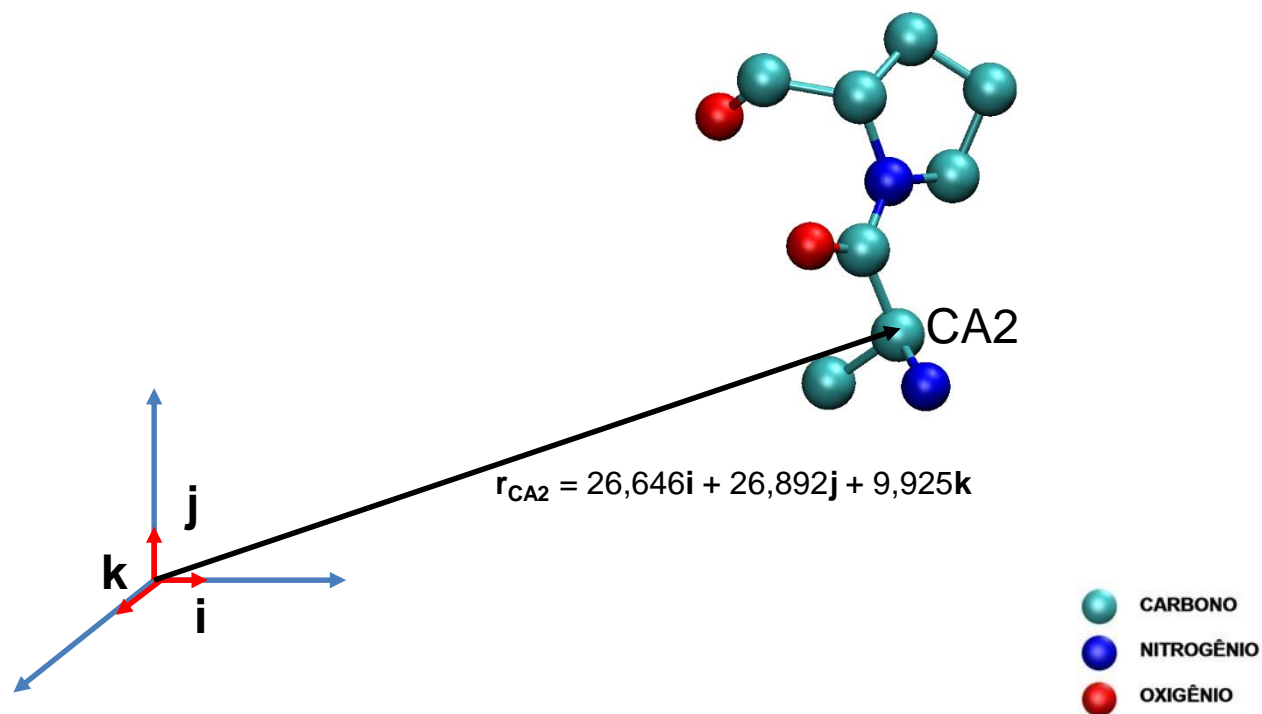


- CARBONO
- NITROGÊNIO
- OXIGÊNIO

As coordenadas atômicas, armazenadas nos arquivos PDB, permitem a determinação direta de diversos parâmetros estruturais, tais como, distâncias, ângulos de ligação e ângulos de torção. Os átomos na estrutura podem ser representados por vetores, onde as coordenadas x,y,z dos átomos são as coordenadas dos vetores, como indicado para o carbono alfa (CA) abaixo. Lembrando-se, vetores apresentam direção, sentido e magnitude (tamanho do vetor). O tamanho do vetor também é chamado de módulo.



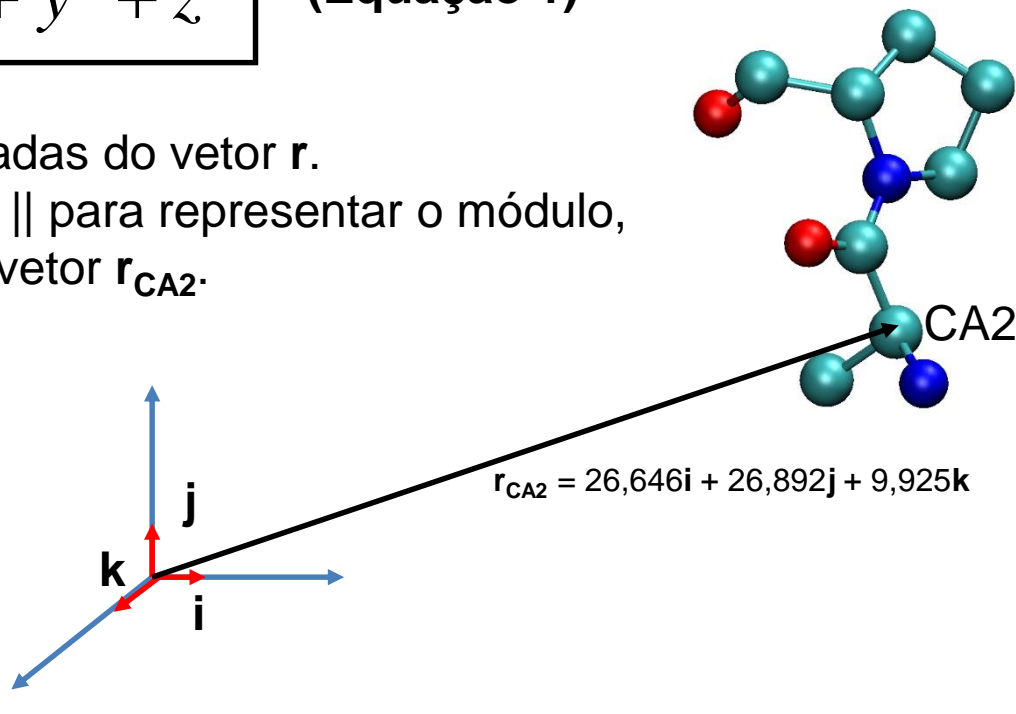
A coordenada ao longo de x é multiplicada pelo vetor unitário \mathbf{i} , a coordenada ao longo de y é multiplicada pelo vetor \mathbf{j} , e a coordenada ao longo z pelo vetor \mathbf{k} . Assim, o vetor $\mathbf{r}_{\text{CA2}} = 26,646\mathbf{i} + 26,892\mathbf{j} + 9,925\mathbf{k}$ indica a posição do átomo carbono alfa da alanina 2 da estrutura 1WE2, como mostrado abaixo. A representação vetorial será usada para determinação dos parâmetros estruturais já citados.



Como destacado, vetores apresentam direção e sentido, indicado pela flecha abaixo. Além disso, apresentam módulo, indicado pelo tamanho da flecha. Para o cálculo do módulo do vetor $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, usamos a seguinte equação:

$$|\mathbf{r}| = \sqrt{x^2 + y^2 + z^2} \quad \text{(Equação 1)}$$

Onde x, y, z são as coordenadas do vetor \mathbf{r} . Usamos as barras verticais $||$ para representar o módulo, assim $|\mathbf{r}_{CA2}|$ é o módulo do vetor \mathbf{r}_{CA2} .



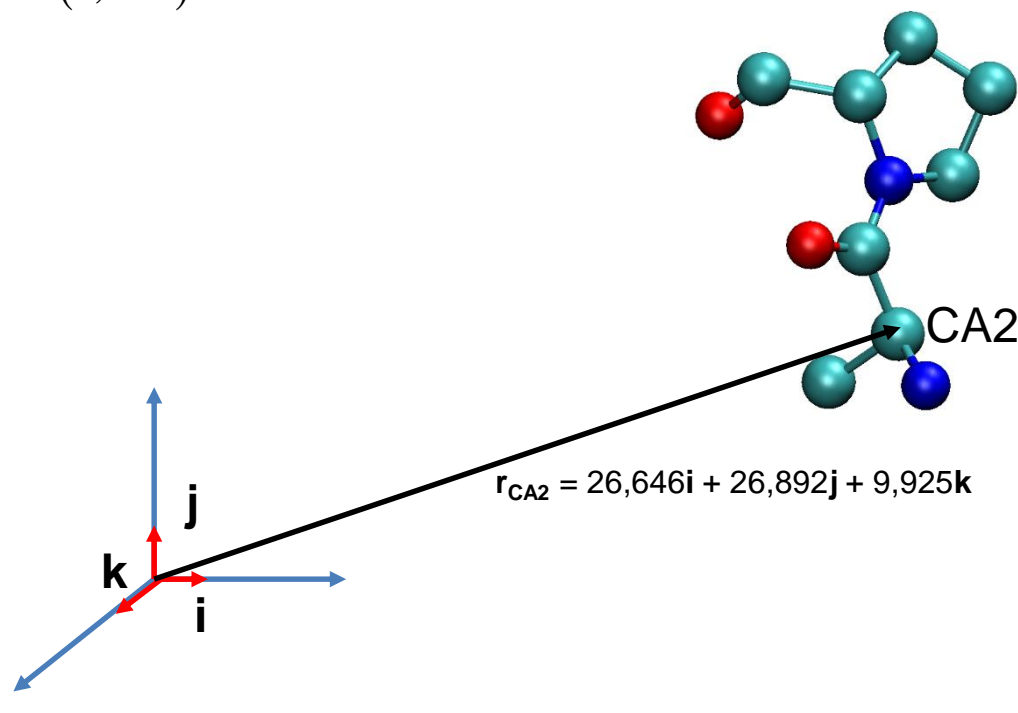
Exemplo 1. Determine o módulo do vetor \mathbf{r}_{CA2} .

Solução

Para determinarmos \mathbf{r}_{CA2} usamos a equação 1, como segue:

$$|\mathbf{r}_{CA2}| = \sqrt{(26,646)^2 + (26,892)^2 + (9,925)^2}$$

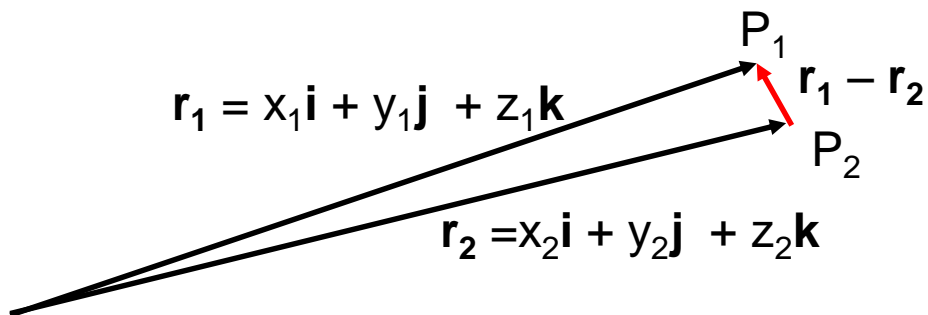
$$|\mathbf{r}_{CA2}| = 39,137 \text{ \AA}$$



Para determinarmos o vetor diferença entre os vetores $\mathbf{r}_1 = x_1\mathbf{i} + y_1\mathbf{j} + z_1\mathbf{k}$ e $\mathbf{r}_2 = x_2\mathbf{i} + y_2\mathbf{j} + z_2\mathbf{k}$, usamos a seguinte equação:

$$\mathbf{r}_1 - \mathbf{r}_2 = (x_1 - x_2)\mathbf{i} + (y_1 - y_2)\mathbf{j} + (z_1 - z_2)\mathbf{k} \quad \text{(Equação 2)}$$

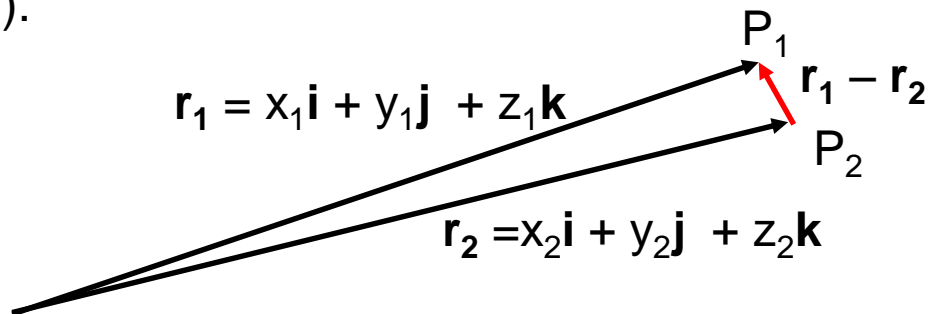
A figura abaixo indica o vetor diferença $\mathbf{r}_1 - \mathbf{r}_2$, entre os pontos P_1 e P_2 .



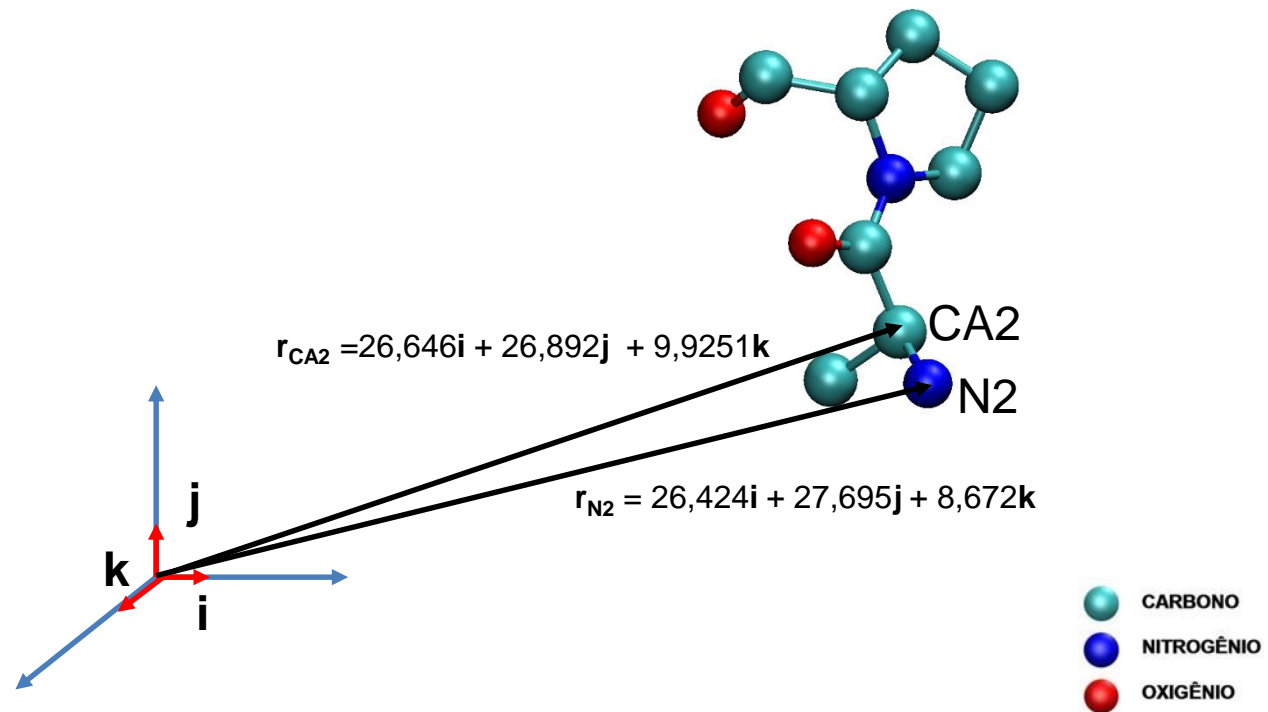
Assim, para determinarmos a distância (d_{12}) entre dois pontos P_1 e P_2 , basta determinarmos o módulo do vetor diferença entre os pontos P_1 e P_2 . Como já temos uma equação para o vetor diferença, basta calcularmos o módulo deste vetor, com a seguinte equação:

$$d_{12} = |\mathbf{r}_1 - \mathbf{r}_2| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} \quad \text{(Equação 3)}$$

Usando-se as coordenadas atômicas de um arquivo PDB, para um par de átomos, teremos a distância em Å (10^{-10} m).



Exemplo2. Determine a distância interatômica para os átomos CA2 e N2, indicados na estrutura abaixo.



Exemplo 2 (Solução). Os vetores são os seguintes:

$$\mathbf{r}_{CA2} = 26,646\mathbf{i} + 26,892\mathbf{j} + 9,9251\mathbf{k}$$

$$\mathbf{r}_{N2} = 26,424\mathbf{i} + 27,695\mathbf{j} + 8,672\mathbf{k}$$

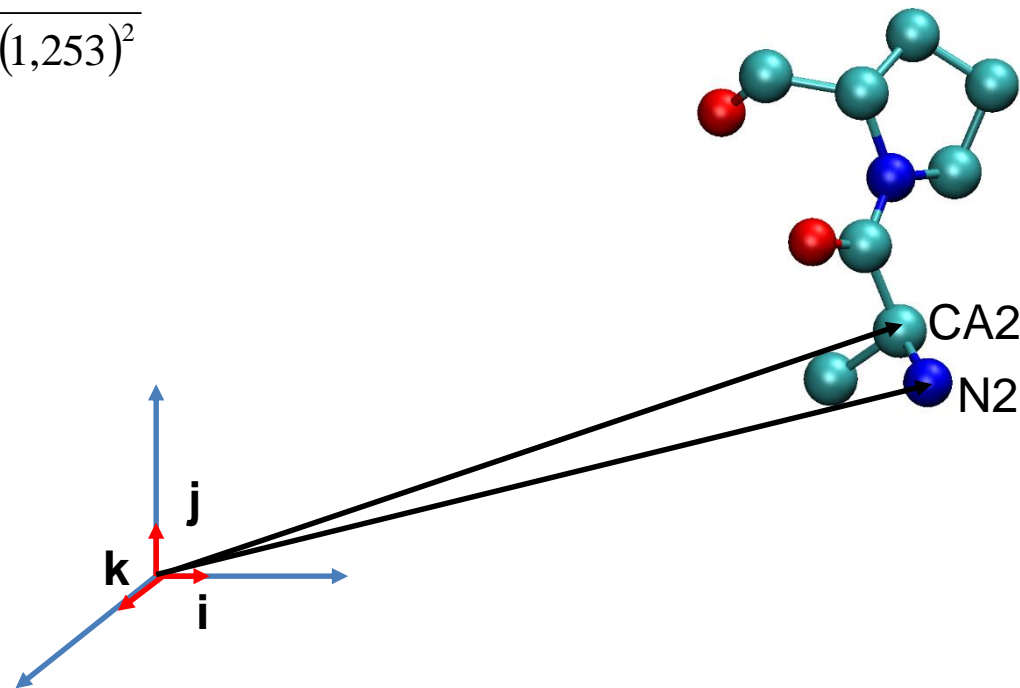
A distância é dada pela equação 3, como segue:

$$d_{CA2,N2} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} = \sqrt{(26,646 - 26,424)^2 + (26,892 - 27,695)^2 + (9,925 - 8,672)^2}$$

$$d_{CA2,N2} = \sqrt{(0,222)^2 + (-0,803)^2 + (1,253)^2}$$

$$d_{CA2,N2} = \sqrt{2.264102}$$

$$d_{CA2,N2} = 1,504 \text{ \AA}$$

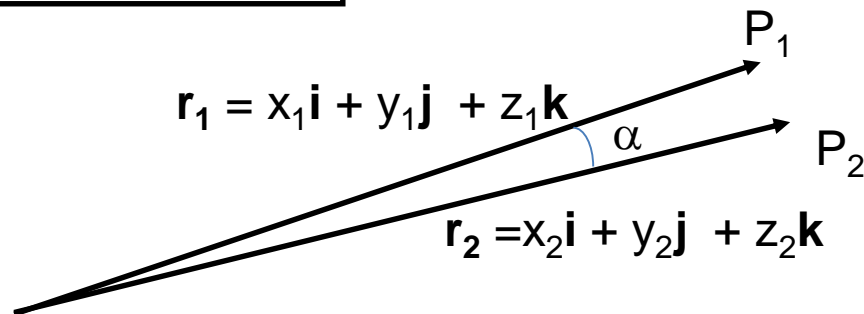


Definimos produto escalar (.) entre os vetores $\mathbf{r}_1 = x_1\mathbf{i} + y_1\mathbf{j} + z_1\mathbf{k}$ e $\mathbf{r}_2 = x_2\mathbf{i} + y_2\mathbf{j} + z_2\mathbf{k}$, pela seguinte equação:

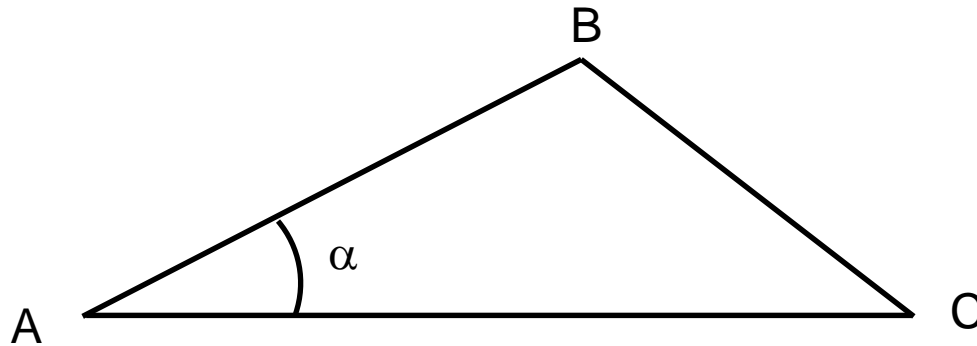
$$\mathbf{r}_1 \cdot \mathbf{r}_2 = x_1x_2 + y_1y_2 + z_1z_2 \quad \text{(Equação 4)}$$

A partir do ângulo α entre os vetores \mathbf{r}_1 e \mathbf{r}_2 , podemos usar a seguinte equação para determinarmos o produto escalar (.):

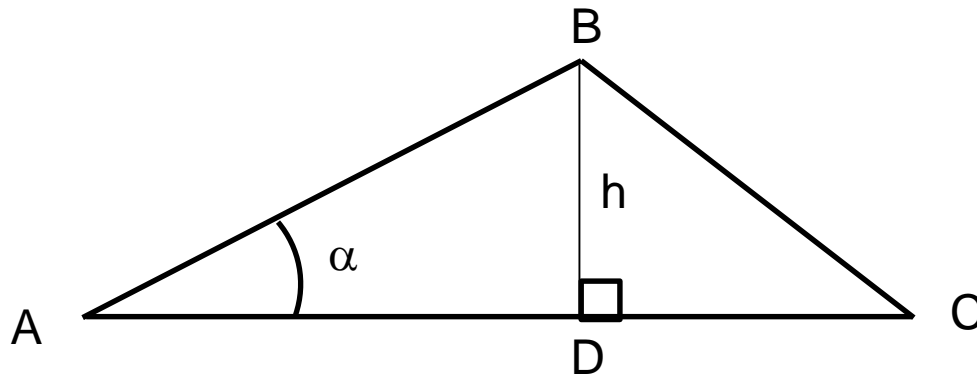
$$\mathbf{r}_1 \cdot \mathbf{r}_2 = |\mathbf{r}_1||\mathbf{r}_2|\cos\alpha \quad \text{(Equação 5)}$$



Para calcularmos o ângulo α de um sistema com três pontos, temos que ter as distâncias entre os pontos. A relação matemática que usaremos é a **lei dos cossenos**. Para revisar o conceito, consideremos o triângulo ABC qualquer abaixo.



O triângulo não é retângulo, mas se traçarmos uma perpendicular à linha AC, a partir do vértice B, temos dois triângulos retângulos, como mostrado abaixo.



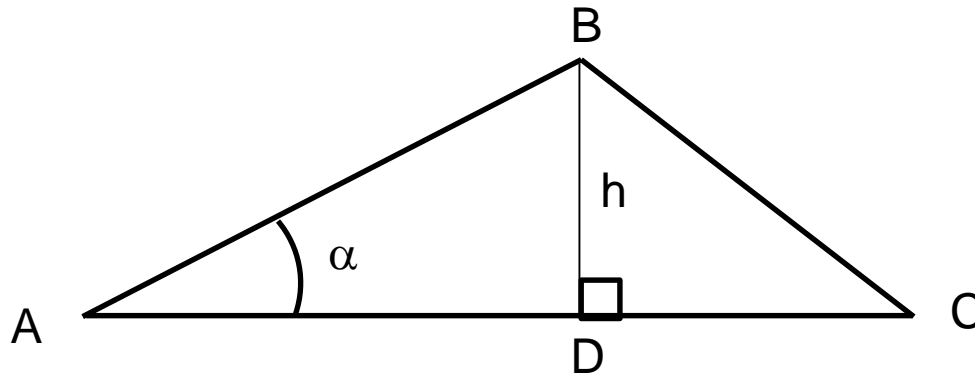
O $\cos \alpha$ é dado pelo cateto adjacente (AD) dividido pela hipotenusa (AB), como segue:

$$\boxed{\cos \alpha = \frac{(AD)}{(AB)}} \quad \text{(Equação 6)}$$

Se usarmos os dois triângulos internos e o teorema de Pitágoras, podemos determinar (AD) em função dos lados do triângulo externo ABC. Considere o triângulo ABD, temos pelo teorema de Pitágoras:

$$(AB)^2 = x^2 + h^2 \quad \Rightarrow \quad \boxed{h^2 = (AB)^2 - x^2} \quad \text{(Equação 7)}$$

$$\text{Do triângulo BCD temos: } (BC)^2 = (CD)^2 + h^2 \Rightarrow \boxed{(BC)^2 = ((AC) - x)^2 + h^2} \quad \text{(Equação 8)}$$



$$\begin{aligned} (BD) &= h \\ (AD) &= x \\ (CD) &= (AC) - x \end{aligned}$$

Só lembrando onde queremos chegar, procuramos uma equação para o $\cos \alpha$ em função dos lados do triângulo ABC. Assim, a equação 6 tem a seguinte forma:

$$\cos \alpha = \frac{(AD)}{(AB)} = \frac{x}{(AB)} \quad \text{(Equação 6)}$$

Substituindo-se a equação 7 na equação 8, chegaremos a uma expressão para x , como segue:

$$(BC)^2 = ((AC) - x)^2 + h^2 \quad \text{(Equação 8)}$$

$$h^2 = (AB)^2 - x^2 \quad \text{(Equação 7)}$$

Assim temos:

$$(BC)^2 = ((AC) - x)^2 + h^2 = (AC)^2 - 2x(AC) + \cancel{x^2} + (AB)^2 - \cancel{x^2}$$

$$(BC)^2 = (AC)^2 - 2x(AC) + (AB)^2 \Rightarrow 2x(AC) = (AC)^2 + (AB)^2 - (BC)^2$$

Isolando-se x , temos:

$$x = \frac{(AB)^2 + (AC)^2 - (BC)^2}{2(AC)} \quad \text{(Equação 9)}$$

Resumindo-se, chegamos a duas equações:

$$\cos\alpha = \frac{(AD)}{(AB)} = \frac{x}{(AB)} \quad \text{(Equação 6)}$$

$$x = \frac{(AB)^2 + (AC)^2 - (BC)^2}{2(AC)} \quad \text{(Equação 9)}$$

Substituindo-se a equação 9 na equação 6, chegamos a:

$$\cos\alpha = \frac{(AD)}{(AB)} = \frac{\left(\frac{(AB)^2 + (AC)^2 - (BC)^2}{2(AC)}\right)}{2(AB)(AC)} \quad \text{(Equação 10)}$$

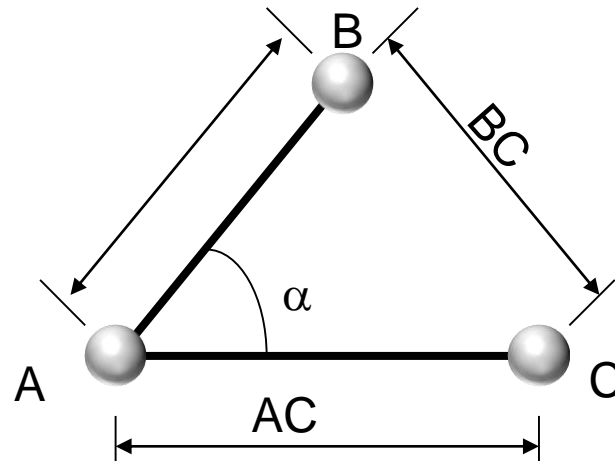
A equação 10 traz o cosseno do ângulo α em função dos lados do triângulo ABC. Nosso objetivo é obtermos o ângulo α , em função da informação sobre os tamanhos dos lados do triângulo ABC ou, de forma equivalente, as distâncias entre os pontos dos vértices do triângulo ABC. Assim faremos uso da função inversa do cosseno, a função arco-cosseno, como segue.

$$\alpha = \arccos \left[\frac{(AB)^2 + (AC)^2 - (BC)^2}{2(AB)(AC)} \right] \quad \text{(Equação 11)}$$

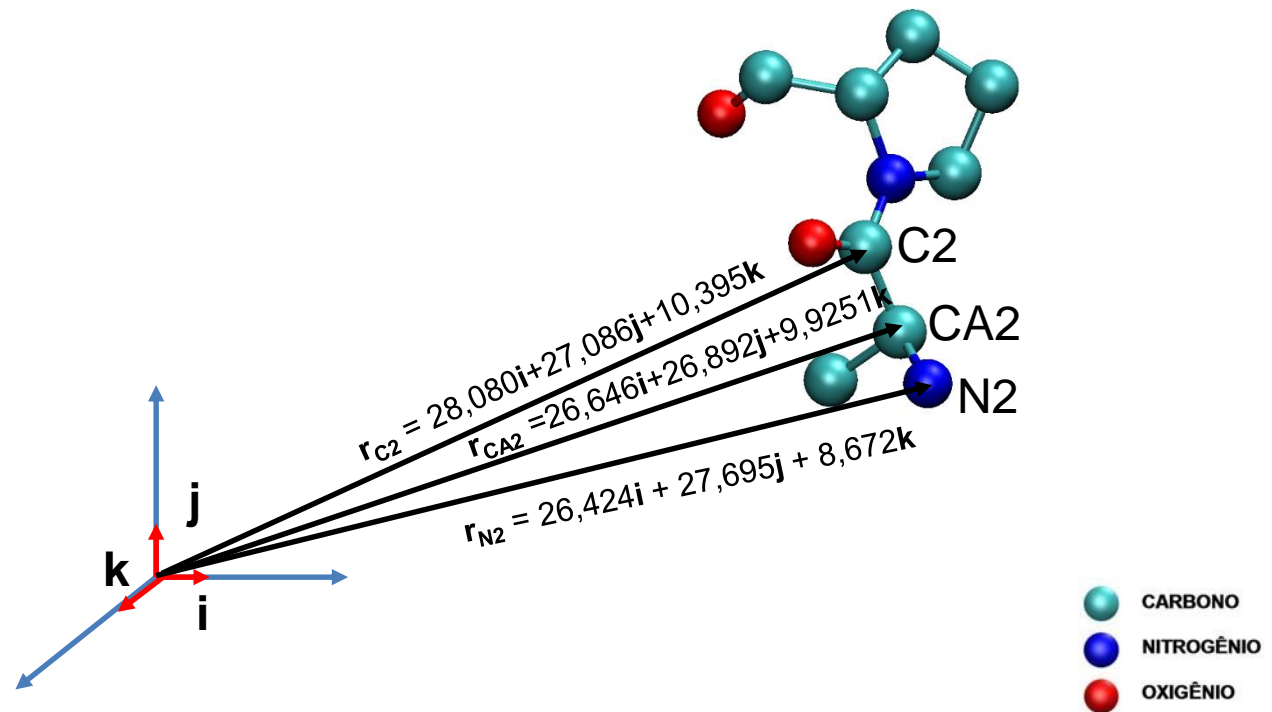
O ângulo de ligação (α), entre as ligações AB e AC, mostrado na figura abaixo, pode ser determinado a partir da seguinte equação:

$$\alpha = \arccos \left[\frac{(AB)^2 + (AC)^2 - (BC)^2}{2(AB)(AC)} \right] \quad \text{(Equação 11)}$$

Onde AB, AC e BC são as distâncias entre os pontos, ou no caso molecular, as distâncias interatômicas, determinadas usando-se a equação da distância interatômica (equação 3), descrita anteriormente.



Exemplo 3. Determine o ângulo de ligação α formados pelos átomos N2 CA2 C2.



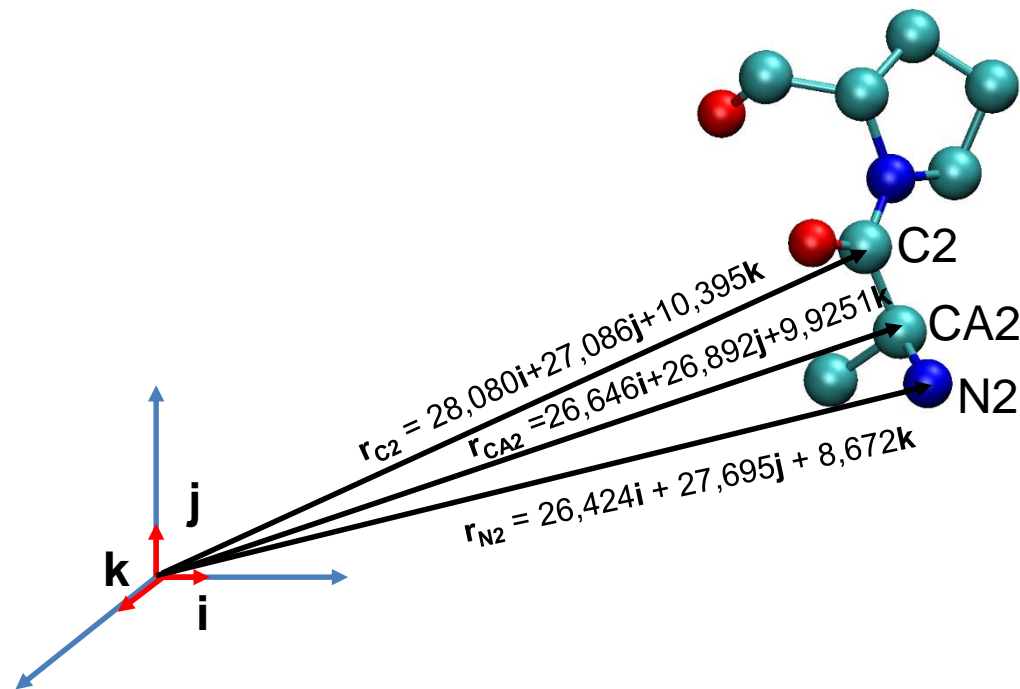
Exemplo 3 (Solução).

Passo 1: Inicialmente usamos a equação 3 para determinar as seguintes distâncias interatômicas: $d_{C2,CA2}$, $d_{CA2,N2}$ e $d_{C2,N2}$. O cálculo é feito a partir da aplicação direta da equação 3, como vimos no exemplo 2. Por isso não iremos repeti-lo aqui. Usaremos os valores das distâncias para os próximos passos.

$$d_{C2,CA2} = 1,520 \text{ \AA}$$

$$d_{CA2,N2} = 1,504 \text{ \AA}$$

$$d_{C2,N2} = 2,470 \text{ \AA}$$



Exemplo 3 (Solução).

Passo 2: Usamos as distâncias interatômicas e substituímos na equação 11, como segue:

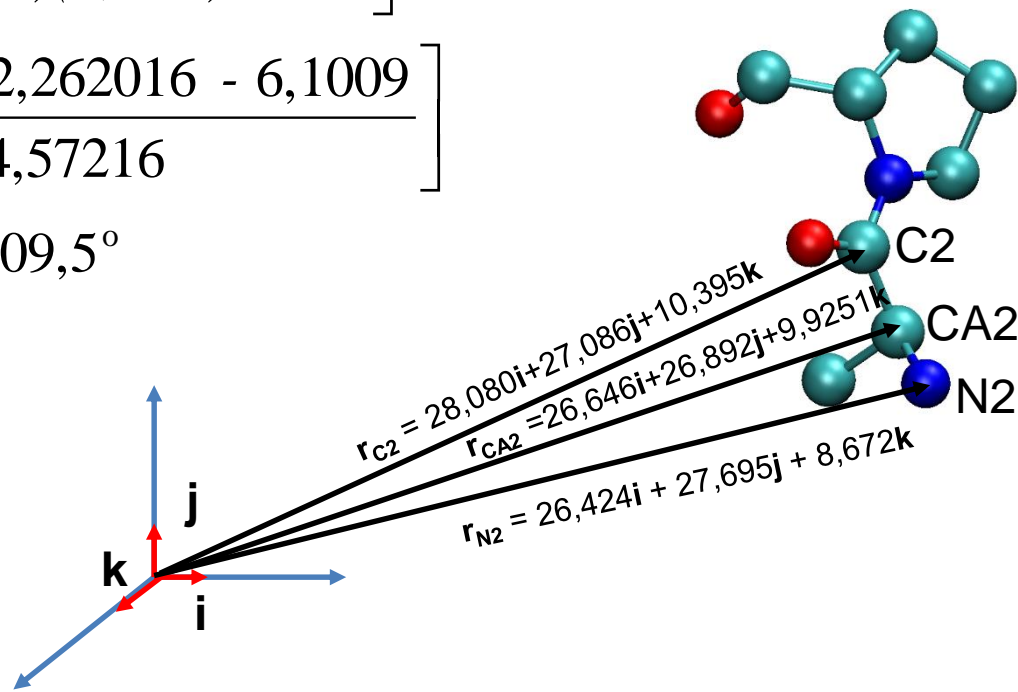
$$\alpha = \arccos \left[\frac{(1,52)^2 + (1,504)^2 - (2,47)^2}{2(1,52)(1,504)} \right]$$

$$\alpha = \arccos \left[\frac{2,3104 + 2,262016 - 6,1009}{4,57216} \right]$$

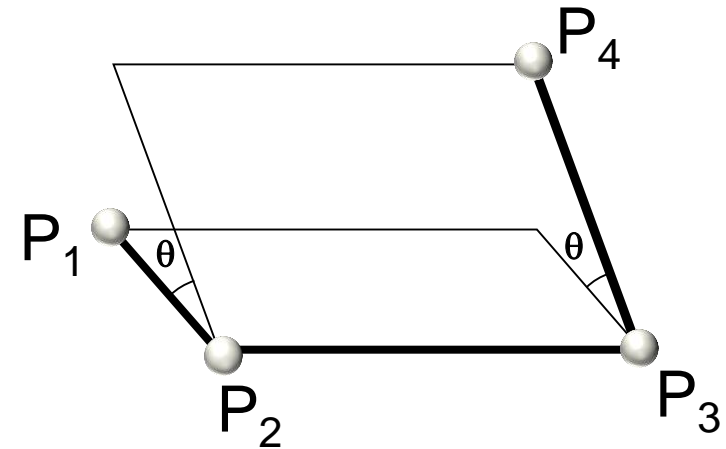
$$\alpha = \arccos [-0,3343] = 109,5^\circ$$

$$\alpha = 109,5^\circ$$

O ângulo de $109,5^\circ$ era o esperado para um carbono tetraédrico, como o carbono alfa (CA).



O ângulo de torção é medido considerando-se duas ternas de pontos, como no sistema de quatro pontos mostrado ao lado, indicados por P_1 , P_2 , P_3 e P_4 . Os pontos P_1 , P_2 e P_3 definem um plano, e os pontos P_2 , P_3 e P_4 um segundo plano. O ângulo de torção envolvendo os pontos P_2 e P_3 é definido como o ângulo formado entre os planos P_1, P_2 e P_3 e P_2, P_3 e P_4 .



Para entendermos a equação do ângulo de torção (θ), necessitamos de alguns conceitos básicos de geometria analítica. Usaremos o conceito de vetor, para determinarmos a interação entre os pontos no sistema. Manteremos a notação de usarmos letras em negrito para representar vetores. Assim, os vetores que indicam as posições dos pontos P_1 , P_2 , P_3 e P_4 são dados por:

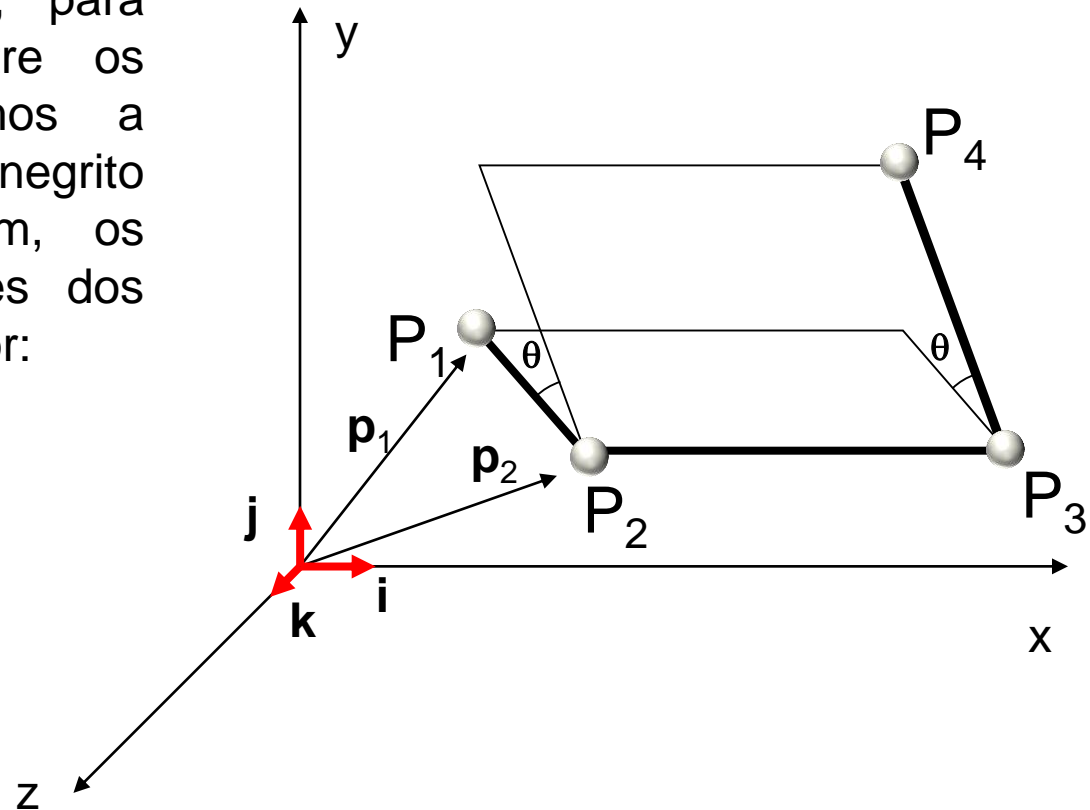
$$\mathbf{p}_1 = x_1\mathbf{i} + y_1\mathbf{j} + z_1\mathbf{k}$$

$$\mathbf{p}_2 = x_2\mathbf{i} + y_2\mathbf{j} + z_2\mathbf{k}$$

$$\mathbf{p}_3 = x_3\mathbf{i} + y_3\mathbf{j} + z_3\mathbf{k}$$

$$\mathbf{p}_4 = x_4\mathbf{i} + y_4\mathbf{j} + z_4\mathbf{k}$$

onde \mathbf{i} , \mathbf{j} e \mathbf{k} são vetores unitários (tamanho (módulo) igual 1) ao longo das direções x , y e z .



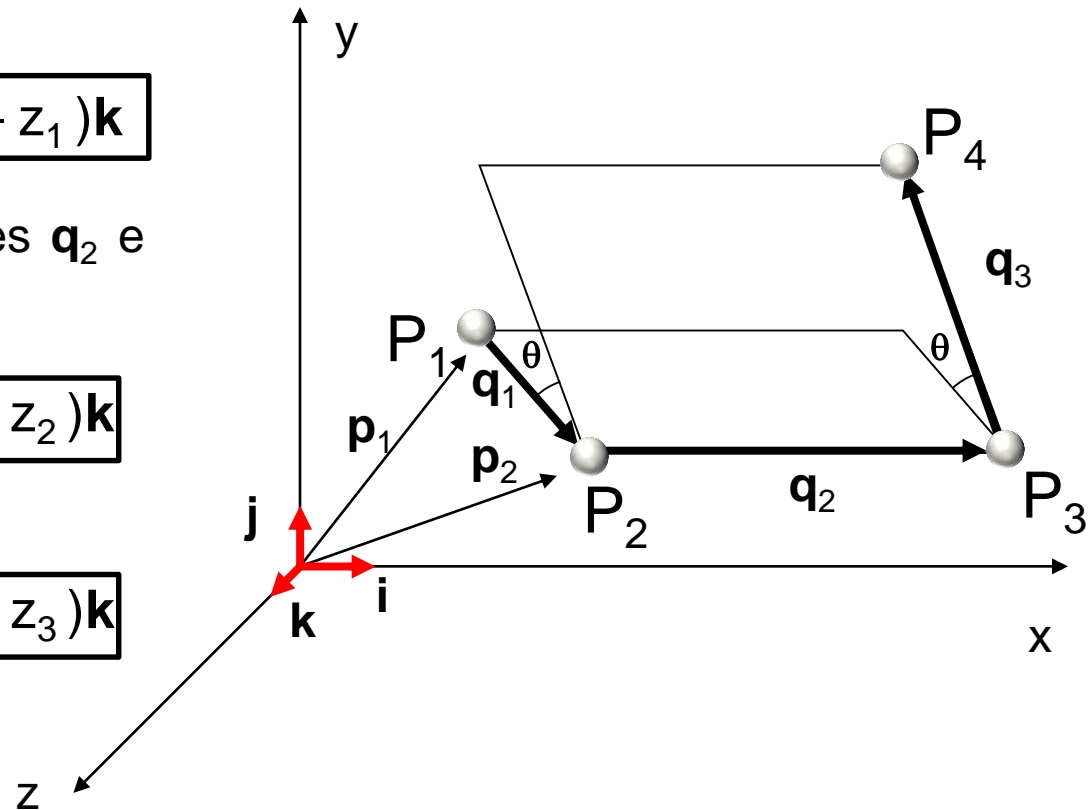
O vetor entre os pontos P_1 e P_2 , aqui chamado de vetor \mathbf{q}_1 , é o vetor subtração, $\mathbf{p}_2 - \mathbf{p}_1$. Em coordenadas cartesianas tem a seguinte expressão:

$$\mathbf{q}_1 = (x_2 - x_1)\mathbf{i} + (y_2 - y_1)\mathbf{j} + (z_2 - z_1)\mathbf{k}$$

De forma análoga temos os vetores \mathbf{q}_2 e \mathbf{q}_3 , como segue:

$$\mathbf{q}_2 = (x_3 - x_2)\mathbf{i} + (y_3 - y_2)\mathbf{j} + (z_3 - z_2)\mathbf{k}$$

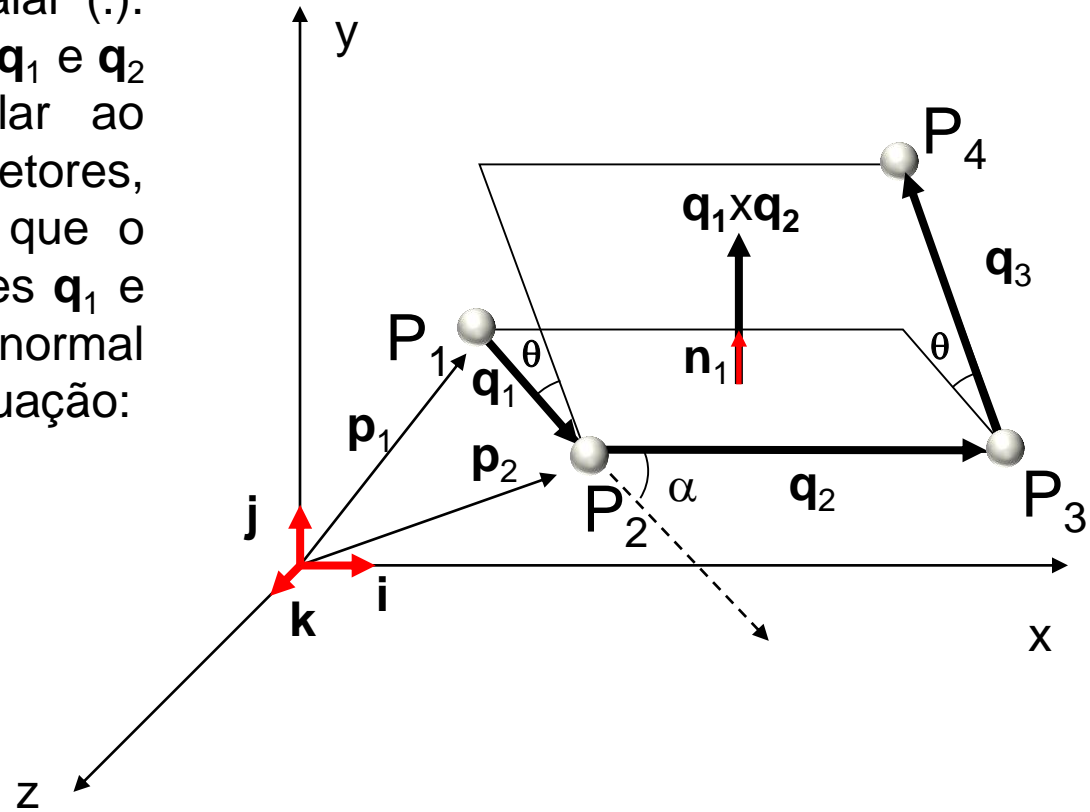
$$\mathbf{q}_3 = (x_4 - x_3)\mathbf{i} + (y_4 - y_3)\mathbf{j} + (z_4 - z_3)\mathbf{k}$$



Para determinarmos o ângulo de torção, precisamos de mais duas definições relacionadas aos vetores, são elas, produto vetorial (\times) e produto escalar (\cdot). O produto vetorial entre os vetores \mathbf{q}_1 e \mathbf{q}_2 é um terceiro vetor perpendicular ao plano, definido por estes dois vetores, assim, na figura ao lado, temos que o vetor \mathbf{n}_1 é perpendicular aos vetores \mathbf{q}_1 e \mathbf{q}_2 e \cdot . O vetor \mathbf{n}_1 é chamado vetor normal ao plano e é dado pela seguinte equação:

$$\mathbf{n}_1 = \frac{\mathbf{q}_1 \times \mathbf{q}_2}{|\mathbf{q}_1 \times \mathbf{q}_2|}$$

$$|\mathbf{q}_1 \times \mathbf{q}_2| = |\mathbf{q}_1| |\mathbf{q}_2| \text{sen } \alpha$$



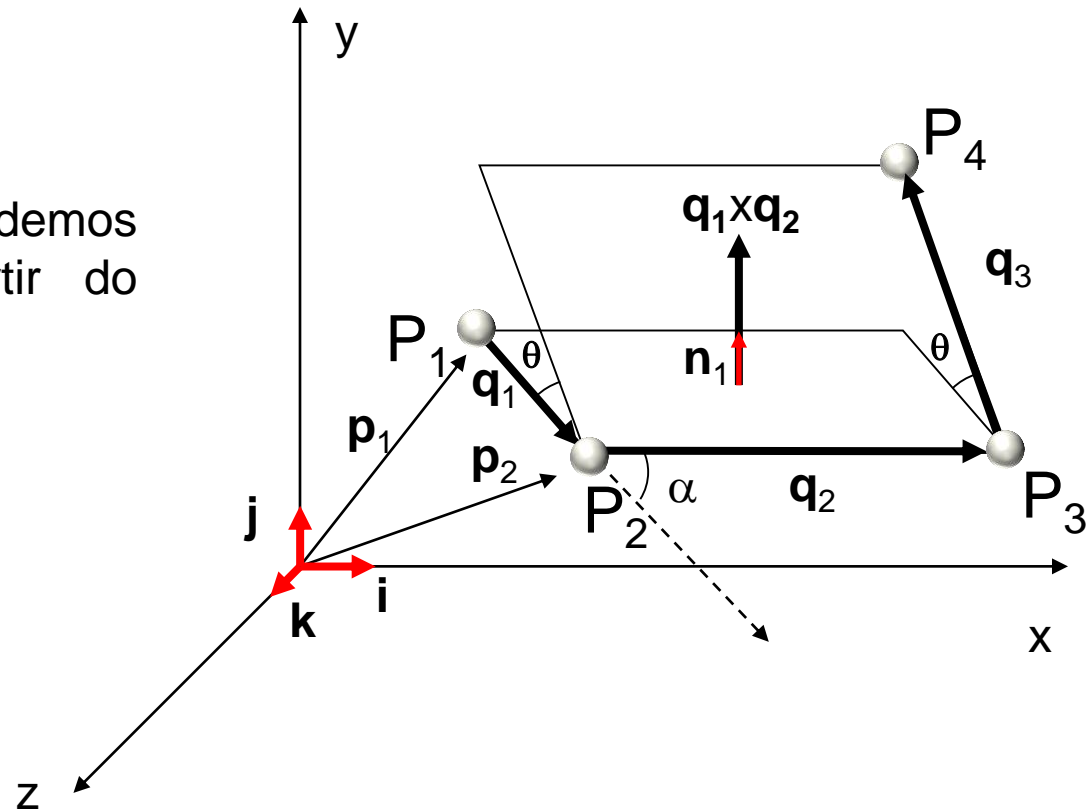
as barras em volta dos vetores indicam seu módulo, ou seja, seu tamanho.

Uma forma alternativa de representarmos o produto vetorial é a forma cartesiana, definida como segue:

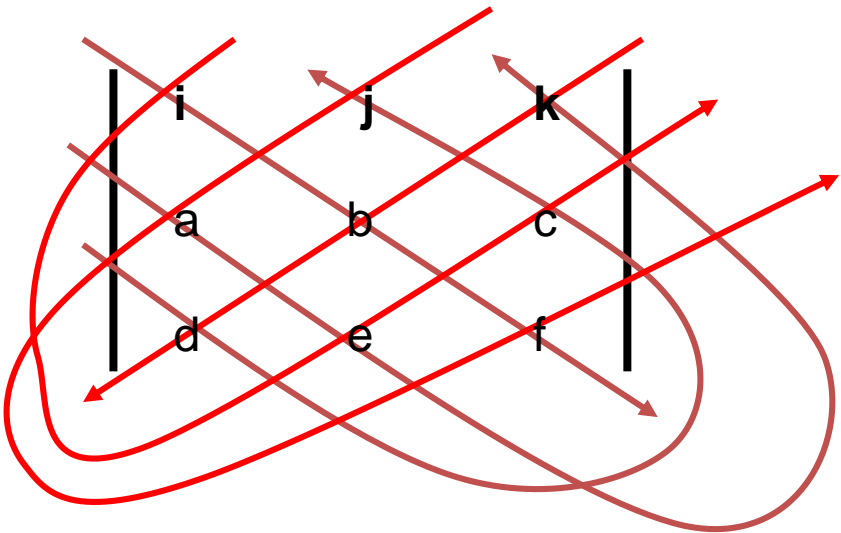
$$\mathbf{q}_1 \times \mathbf{q}_2 = (a\mathbf{i} + b\mathbf{j} + c\mathbf{k}) \times (d\mathbf{i} + e\mathbf{j} + f\mathbf{k})$$

O “x” indica o produto vetorial. Podemos obter o produto vetorial a partir do determinante da matriz abaixo.

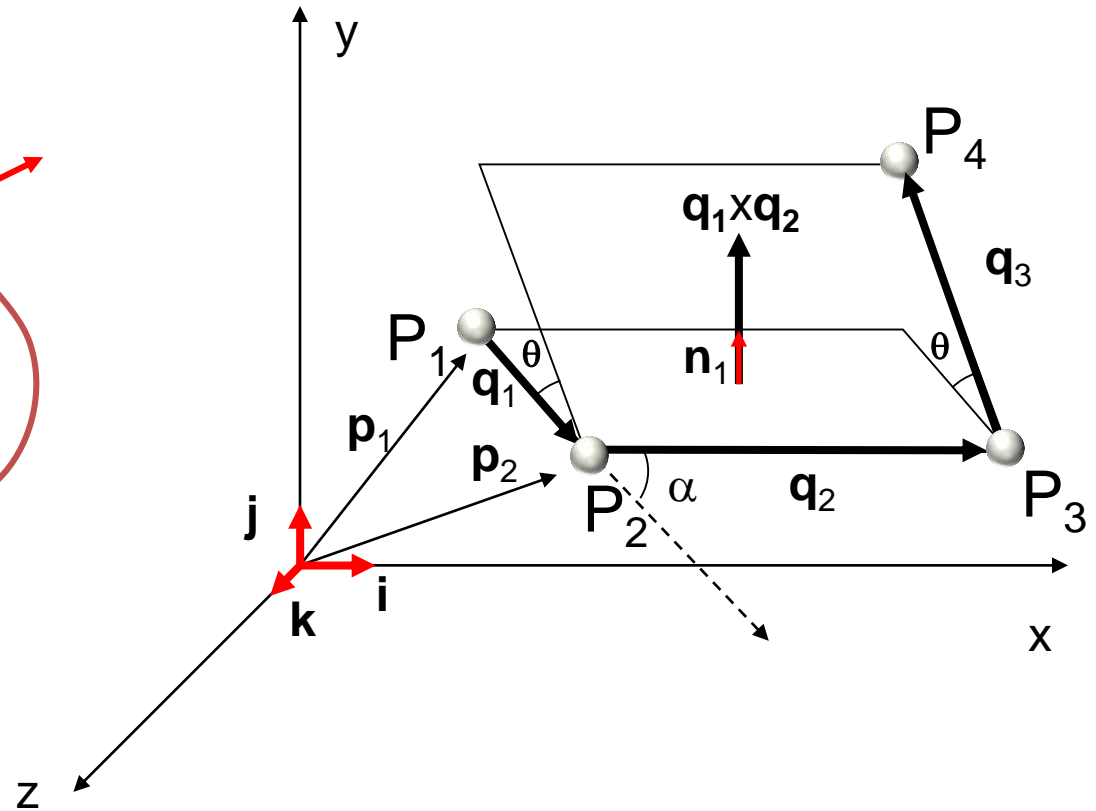
$$\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a & b & c \\ d & e & f \end{vmatrix}$$



Calculamos o determinante da seguinte forma:



$$\begin{aligned} \mathbf{q}_1 \times \mathbf{q}_2 &= \mathbf{i} (b f - e c) \\ &+ \mathbf{j} (d c - a f) \\ &+ \mathbf{k} (a e - b d) \end{aligned}$$



Como já vimos na equação 5, o produto escalar entre dois vetores \mathbf{q}_1 e \mathbf{q}_2 é dado por:

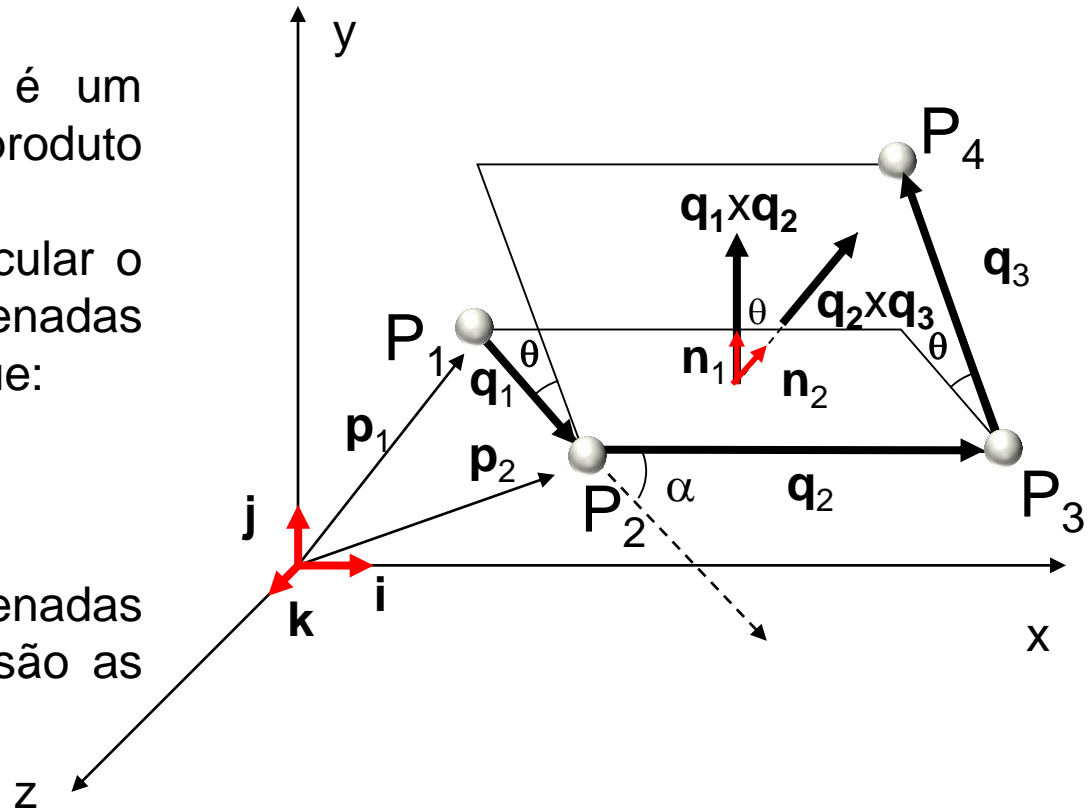
$$\mathbf{q}_1 \cdot \mathbf{q}_2 = |\mathbf{q}_1| |\mathbf{q}_2| \cos \alpha$$

O resultado do produto vetorial é um vetor, enquanto o resultado do produto escalar é um número puro.

Vimos, também, que podemos calcular o produto escalar a partir das coordenadas cartesianas dos vetores, como segue:

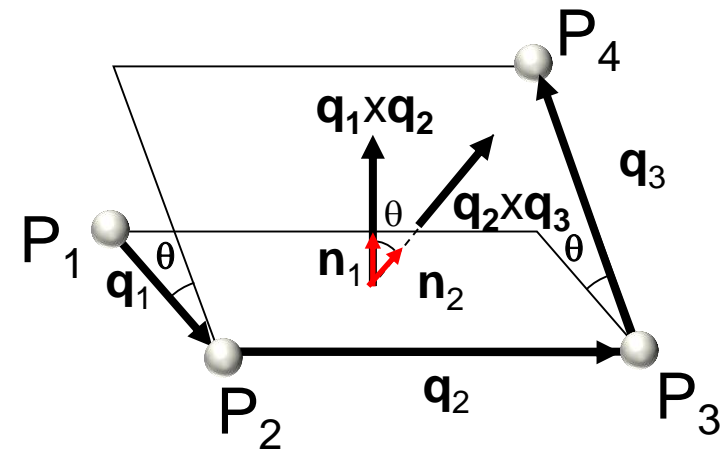
$$\mathbf{q}_1 \cdot \mathbf{q}_2 = a.d + b.e + c.f$$

Onde a , b e c são as coordenadas cartesianas do vetor \mathbf{q}_1 e d , e , f são as coordenadas do vetor \mathbf{q}_2 .



Os programas que calculam ângulos de torção, usam a geometria da figura ao lado para construir uma equação computacionalmente eficiente, baseada na função arco-tangente, especificamente uma implementação desta função é chamada atan2. Bem, para usarmos a função atan2, temos que usar os vetores unitários normais aos planos pelos pontos P_1, P_2, P_3 e P_2, P_3, P_4 . Chamando-se \mathbf{n}_1 o vetor normal ao primeiro plano e \mathbf{n}_2 o vetor normal ao segundo plano temos:

$$\mathbf{n}_1 = \frac{\mathbf{q}_1 \times \mathbf{q}_2}{|\mathbf{q}_1 \times \mathbf{q}_2|} \quad \text{e} \quad \mathbf{n}_2 = \frac{\mathbf{q}_2 \times \mathbf{q}_3}{|\mathbf{q}_2 \times \mathbf{q}_3|}$$



Definimos os vetores unitários pelas equações abaixo:

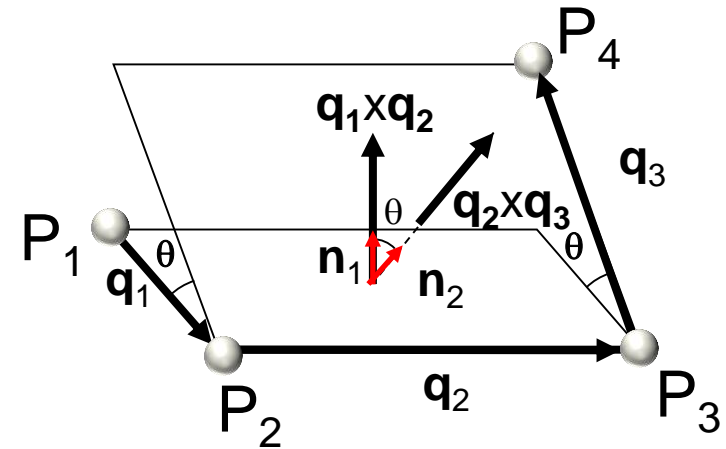
$$\begin{aligned} \mathbf{u}_1 &= \mathbf{n}_2 \\ \mathbf{u}_3 &= \frac{\mathbf{q}_2}{|\mathbf{q}_2|} \\ \mathbf{u}_2 &= \mathbf{u}_3 \times \mathbf{u}_1 \end{aligned}$$

O cosseno e seno de θ são dados por:

$$\begin{aligned} \cos \theta &= \mathbf{n}_1 \cdot \mathbf{u}_1 \\ \sin \theta &= \mathbf{n}_1 \cdot \mathbf{u}_2 \end{aligned}$$

O ângulo θ é dado por

$$\theta = -a \tan 2 \left(\frac{\mathbf{n}_1 \cdot \mathbf{u}_1}{\mathbf{n}_1 \cdot \mathbf{u}_2} \right)$$



Resumindo-se, para o cálculo do ângulo de torção θ de um sistema de quatro pontos, como o mostrado ao lado, temos os seguintes passos:

1) Determinar os vetores \mathbf{q}_1 , \mathbf{q}_2 e \mathbf{q}_3 :

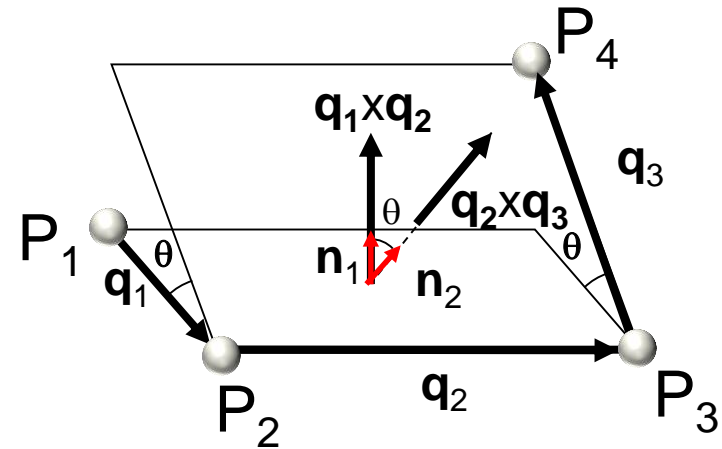
$$\mathbf{q}_1 = (x_2 - x_1)\mathbf{i} + (y_2 - y_1)\mathbf{j} + (z_2 - z_1)\mathbf{k}$$

$$\mathbf{q}_2 = (x_3 - x_2)\mathbf{i} + (y_3 - y_2)\mathbf{j} + (z_3 - z_2)\mathbf{k}$$

$$\mathbf{q}_3 = (x_4 - x_3)\mathbf{i} + (y_4 - y_3)\mathbf{j} + (z_4 - z_3)\mathbf{k}$$

2) Calcular os produtos vetoriais $\mathbf{q}_1 \times \mathbf{q}_2$ e $\mathbf{q}_2 \times \mathbf{q}_3$.

$$\begin{array}{l} \mathbf{q}_1 \times \mathbf{q}_2 \\ \mathbf{q}_2 \times \mathbf{q}_3 \end{array}$$



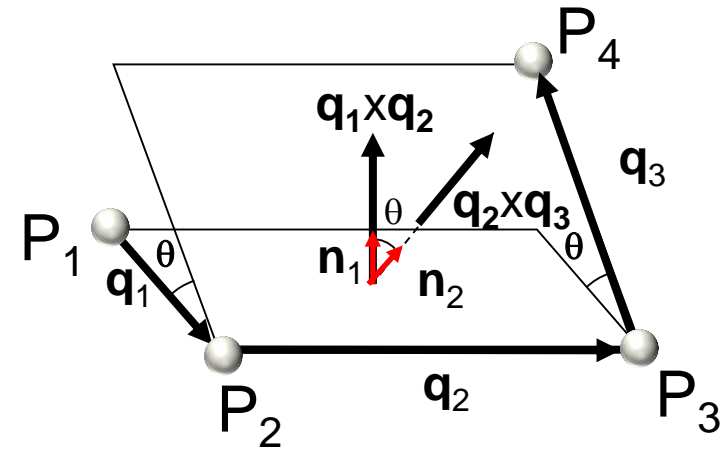
3) Calcular as normais aos planos:

$$\mathbf{n}_1 = \frac{\mathbf{q}_1 \times \mathbf{q}_2}{|\mathbf{q}_1 \times \mathbf{q}_2|}$$

$$\mathbf{n}_2 = \frac{\mathbf{q}_2 \times \mathbf{q}_3}{|\mathbf{q}_2 \times \mathbf{q}_3|}$$

4) Calcular os vetores unitários ortogonais:

$$\begin{aligned} \mathbf{u}_1 &= \mathbf{n}_2 \\ \mathbf{u}_3 &= \frac{\mathbf{q}_2}{|\mathbf{q}_2|} \\ \mathbf{u}_2 &= \mathbf{u}_3 \times \mathbf{u}_1 \end{aligned}$$



5) Calcular o ângulo de torção θ :

$$\begin{aligned} \cos \theta &= \mathbf{n}_1 \cdot \mathbf{u}_1 \\ \sin \theta &= \mathbf{n}_1 \cdot \mathbf{u}_2 \end{aligned}$$

$$\theta = -a \tan 2 \left(\frac{\mathbf{n}_1 \cdot \mathbf{u}_1}{\mathbf{n}_1 \cdot \mathbf{u}_2} \right)$$

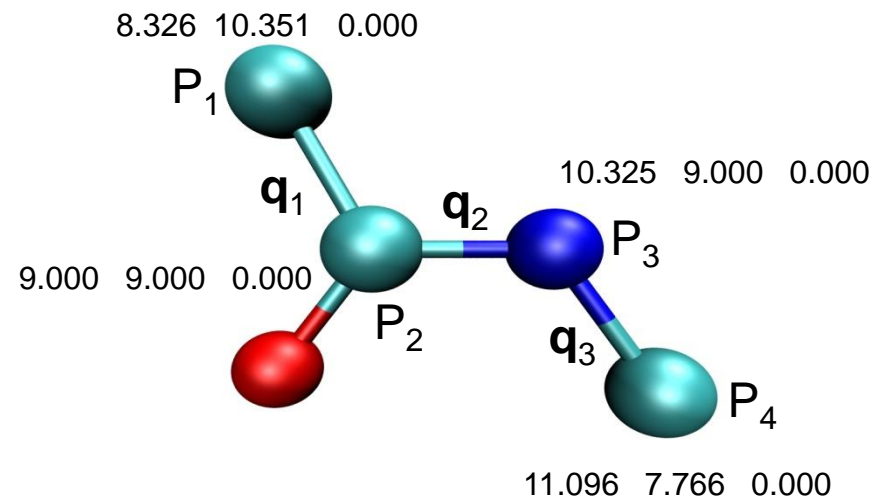
Exemplo 4. Determine o ângulo de torção θ , dos planos formados pelos pontos P_1 , P_2 , P_3 e P_4 , sabendo-se que as coordenadas dos pontos são as seguintes:

$$\mathbf{p}_1 = 8,326\mathbf{i} + 10,351\mathbf{j} + 0,000\mathbf{k}$$

$$\mathbf{p}_2 = 9,000\mathbf{i} + 9,000\mathbf{j} + 0,000\mathbf{k}$$

$$\mathbf{p}_3 = 10,325\mathbf{i} + 9,000\mathbf{j} + 0,000\mathbf{k}$$

$$\mathbf{p}_4 = 11,096\mathbf{i} + 7,766\mathbf{j} + 0,000\mathbf{k}$$



Solução

Passo 1: Para determinar o ângulo de torção, dos planos formados pelos pontos P_1 , P_2 , P_3 e P_4 , precisamos inicialmente calcular os vetores \mathbf{p} , como segue:

$$\mathbf{q}_1 = (x_2 - x_1)\mathbf{i} + (y_2 - y_1)\mathbf{j} + (z_2 - z_1)\mathbf{k} = (0,674)\mathbf{i} + (-1,351)\mathbf{j} = 0,674\mathbf{i} - 1,351\mathbf{j}$$

$$\mathbf{q}_2 = (x_3 - x_2)\mathbf{i} + (y_3 - y_2)\mathbf{j} + (z_3 - z_2)\mathbf{k} = (1,351)\mathbf{i} + (0)\mathbf{j} = 1,351\mathbf{i}$$

$$\mathbf{q}_3 = (x_4 - x_3)\mathbf{i} + (y_4 - y_3)\mathbf{j} + (z_4 - z_3)\mathbf{k} = (0,771)\mathbf{i} + (-1,234)\mathbf{j} = 0,771\mathbf{i} - 1,234\mathbf{j}$$

Solução (Continuação)

Passo 2: Agora com os vetores \mathbf{q} determinados, podemos calcular os produtos vetoriais como segue:

$$\mathbf{q}_1 \times \mathbf{q}_2 = (0,674\mathbf{i} - 1,351\mathbf{j}) \times (1,351\mathbf{i}) = 1,8252\mathbf{k}$$

$$\mathbf{q}_2 \times \mathbf{q}_3 = (1,351\mathbf{i}) \times (0,771\mathbf{i} - 1,234\mathbf{j}) = -1,6671\mathbf{k}$$

Passo 3: Calcularemos as normais aos planos:

$$\mathbf{n}_1 = \frac{\mathbf{q}_1 \times \mathbf{q}_2}{|\mathbf{q}_1 \times \mathbf{q}_2|} = \frac{1,8252\mathbf{k}}{1,8252} = \mathbf{k}$$

$$\mathbf{n}_2 = \frac{\mathbf{q}_2 \times \mathbf{q}_3}{|\mathbf{q}_2 \times \mathbf{q}_3|} = \frac{-1,6671\mathbf{k}}{1,6671} = -\mathbf{k}$$

Regra do produto vetorial:

$\mathbf{i} \times \mathbf{i} = 0$	$\mathbf{j} \times \mathbf{j} = 0$	$\mathbf{k} \times \mathbf{k} = 0$
$\mathbf{i} \times \mathbf{j} = \mathbf{k}$	$\mathbf{j} \times \mathbf{k} = \mathbf{i}$	$\mathbf{k} \times \mathbf{i} = \mathbf{j}$
$\mathbf{i} \times \mathbf{k} = -\mathbf{j}$	$\mathbf{j} \times \mathbf{i} = -\mathbf{k}$	$\mathbf{k} \times \mathbf{j} = -\mathbf{i}$

Solução (Continuação)

Passo 4: Calcularemos os vetores unitários:

$$\mathbf{u}_1 = \mathbf{n}_2 = -\mathbf{k}$$

$$\mathbf{u}_3 = \frac{\mathbf{q}_2}{|\mathbf{q}_2|} = \frac{1,351\mathbf{i}}{1,351} = \mathbf{i}$$

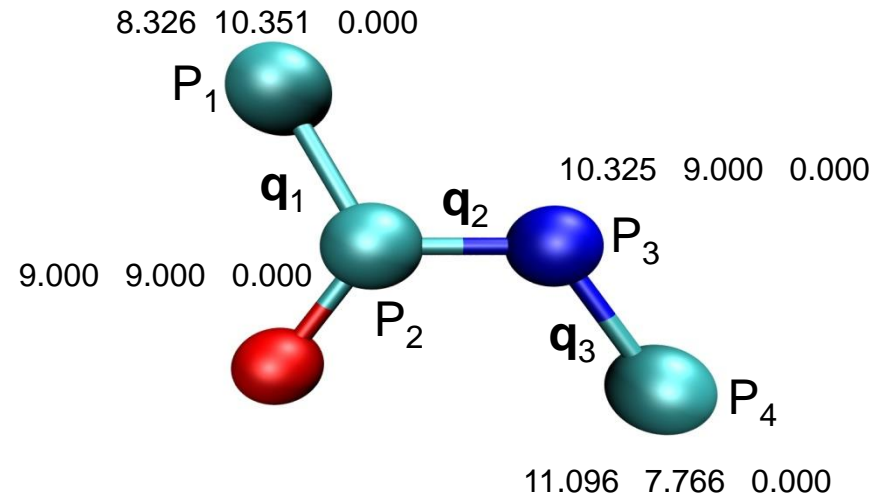
$$\mathbf{u}_2 = \mathbf{u}_3 \times \mathbf{u}_1 = \mathbf{i} \times (-\mathbf{k}) = \mathbf{j}$$

Passo 5: Finalmente o ângulo de torção:

$$\cos -\theta = \mathbf{n}_1 \cdot \mathbf{u}_1 = \mathbf{k} \cdot (-\mathbf{k}) = -1$$

$$\text{sen} -\theta = \mathbf{n}_1 \cdot \mathbf{u}_2 = \mathbf{k} \cdot \mathbf{j} = 1$$

$$\theta = -\arctan 2\left(\frac{1}{-1}\right) = -180^\circ$$



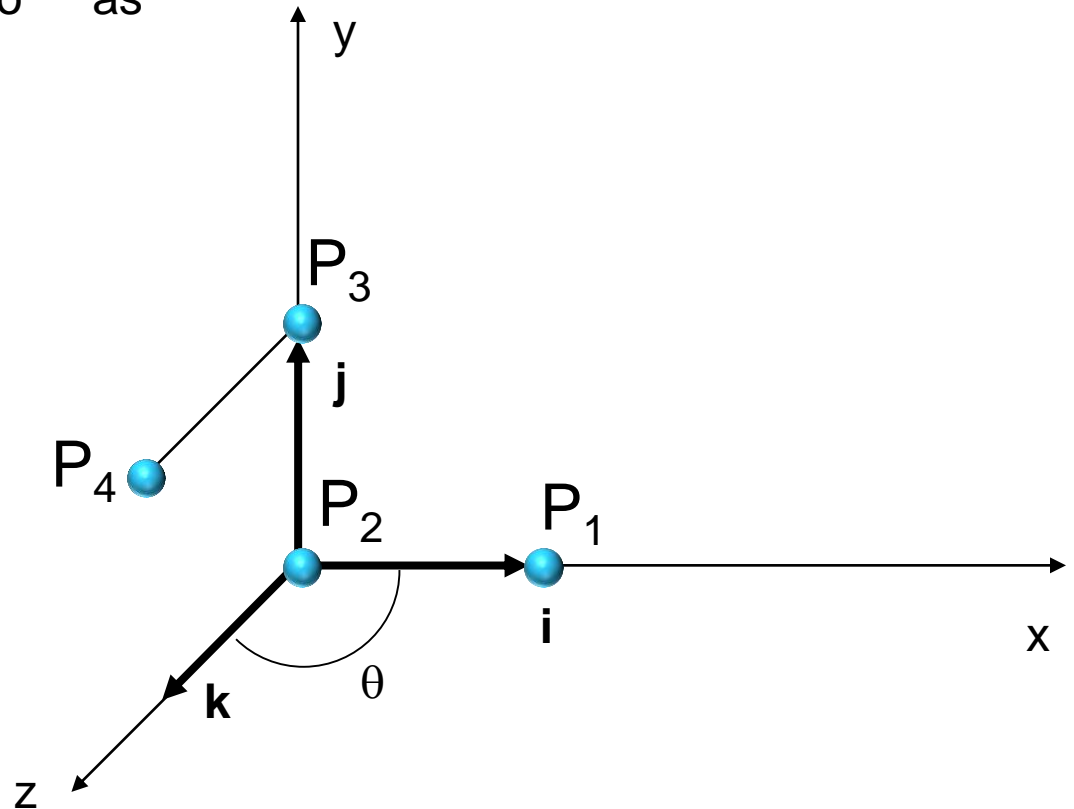
Exemplo 5: Determine o ângulo de torção θ , dos planos formados pelos pontos P_1 , P_2 , P_3 e P_4 , sabendo-se que as coordenadas dos pontos são as seguintes:

$$\mathbf{p}_1 = \mathbf{i}$$

$$\mathbf{p}_2 = \mathbf{0}$$

$$\mathbf{p}_3 = \mathbf{j}$$

$$\mathbf{p}_4 = \mathbf{j} + \mathbf{k}$$



Solução (Continuação)

Passo 1: Para determinar o ângulo de torção, dos planos formados pelos pontos P_1 , P_2 , P_3 e P_4 , precisamos inicialmente calcular os vetores \mathbf{p} , como segue:

$$\mathbf{q}_1 = (x_2 - x_1)\mathbf{i} + (y_2 - y_1)\mathbf{j} + (z_2 - z_1)\mathbf{k} = (0 - 1)\mathbf{i} + (0 - 0)\mathbf{j} + (0 - 0)\mathbf{k} = -\mathbf{i}$$

$$\mathbf{q}_2 = (x_3 - x_2)\mathbf{i} + (y_3 - y_2)\mathbf{j} + (z_3 - z_2)\mathbf{k} = (0 - 0)\mathbf{i} + (1 - 0)\mathbf{j} + (0 - 0)\mathbf{k} = \mathbf{j}$$

$$\mathbf{q}_3 = (x_4 - x_3)\mathbf{i} + (y_4 - y_3)\mathbf{j} + (z_4 - z_3)\mathbf{k} = (0 - 0)\mathbf{i} + (1 - 1)\mathbf{j} + (1 - 0)\mathbf{k} = \mathbf{k}$$

Solução (Continuação)

Passo 2: Agora com os vetores \mathbf{q} determinados, podemos calcular os produtos vetoriais, como segue:

$$\mathbf{q}_1 \times \mathbf{q}_2 = -\mathbf{i} \times \mathbf{j} = -\mathbf{k}$$

$$\mathbf{q}_2 \times \mathbf{q}_3 = \mathbf{j} \times \mathbf{k} = \mathbf{i}$$

Regra do produto vetorial:

$\mathbf{i} \times \mathbf{i} = 0$	$\mathbf{j} \times \mathbf{j} = 0$	$\mathbf{k} \times \mathbf{k} = 0$
$\mathbf{i} \times \mathbf{j} = \mathbf{k}$	$\mathbf{j} \times \mathbf{k} = \mathbf{i}$	$\mathbf{k} \times \mathbf{i} = \mathbf{j}$
$\mathbf{i} \times \mathbf{k} = -\mathbf{j}$	$\mathbf{j} \times \mathbf{i} = -\mathbf{k}$	$\mathbf{k} \times \mathbf{j} = -\mathbf{i}$

Passo 3: Calcularemos as normais aos planos:

$$\mathbf{n}_1 = \frac{\mathbf{q}_1 \times \mathbf{q}_2}{|\mathbf{q}_1 \times \mathbf{q}_2|} = \frac{-\mathbf{k}}{1} = -\mathbf{k}$$

$$\mathbf{n}_2 = \frac{\mathbf{q}_2 \times \mathbf{q}_3}{|\mathbf{q}_2 \times \mathbf{q}_3|} = \frac{\mathbf{i}}{1} = \mathbf{i}$$

Solução (Continuação)

Passo 4: Calcularemos os vetores unitários:

$$\mathbf{u}_1 = \mathbf{n}_2 = \mathbf{i}$$

$$\mathbf{u}_3 = \frac{\mathbf{q}_2}{|\mathbf{q}_2|} = \mathbf{j}$$

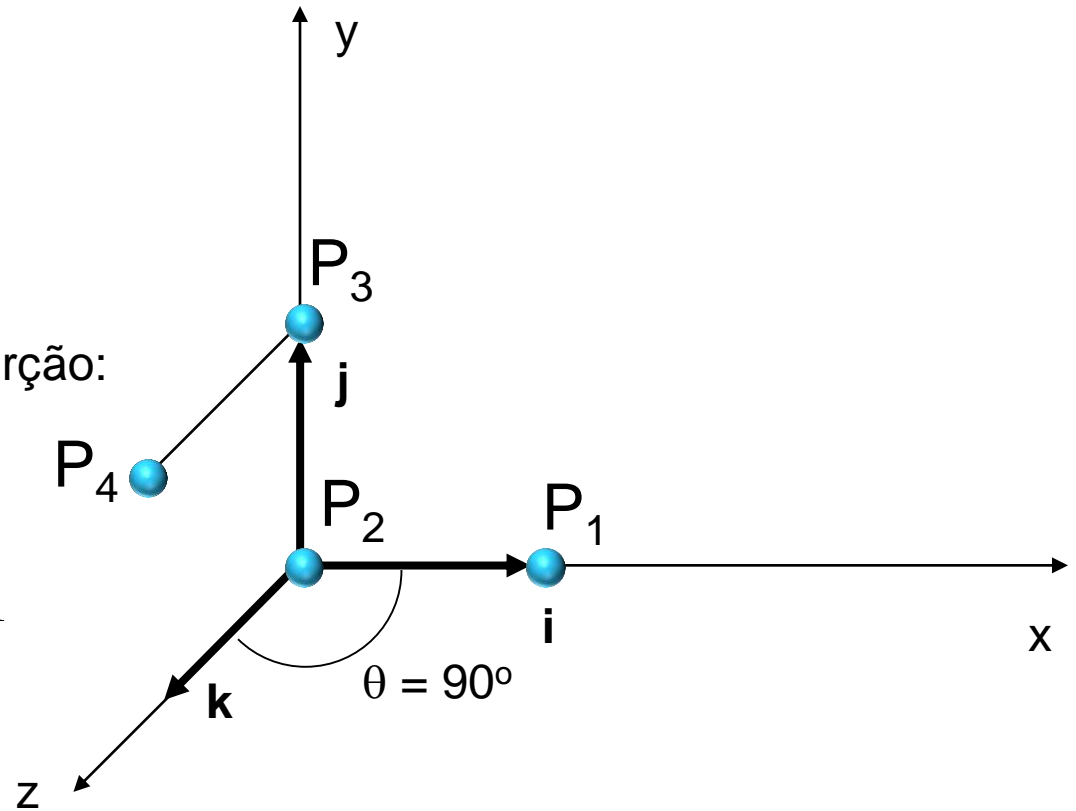
$$\mathbf{u}_2 = \mathbf{u}_3 \times \mathbf{u}_1 = \mathbf{j} \times \mathbf{i} = -\mathbf{k}$$

Passo 5: Finalmente o ângulo de torção:

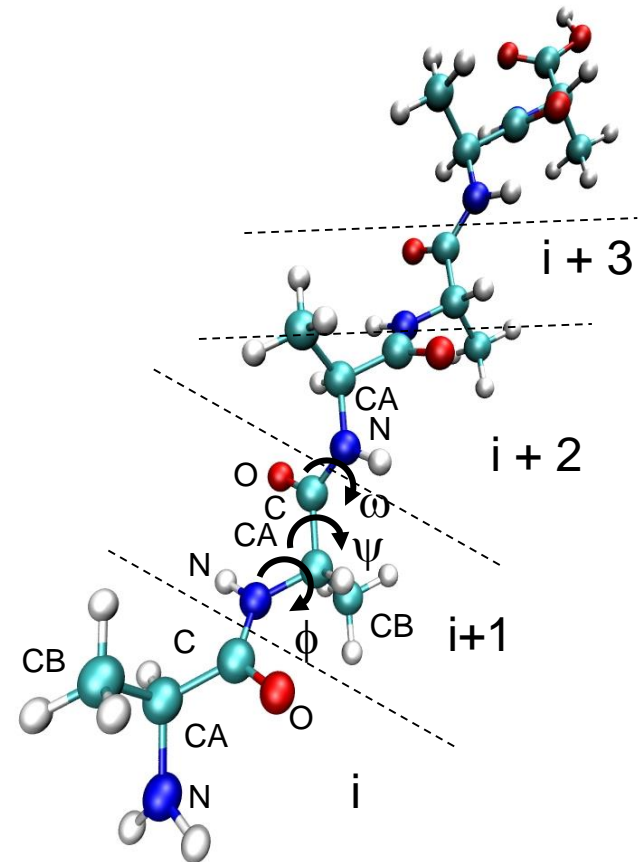
$$\cos -\theta = \mathbf{n}_1 \cdot \mathbf{u}_1 = (-\mathbf{k}) \cdot \mathbf{i} = 0$$

$$\text{sen} -\theta = \mathbf{n}_1 \cdot \mathbf{u}_2 = (-\mathbf{k}) \cdot (-\mathbf{k}) = 1$$

$$\theta = -a \tan 2\left(\frac{1}{0}\right) = -90^\circ$$

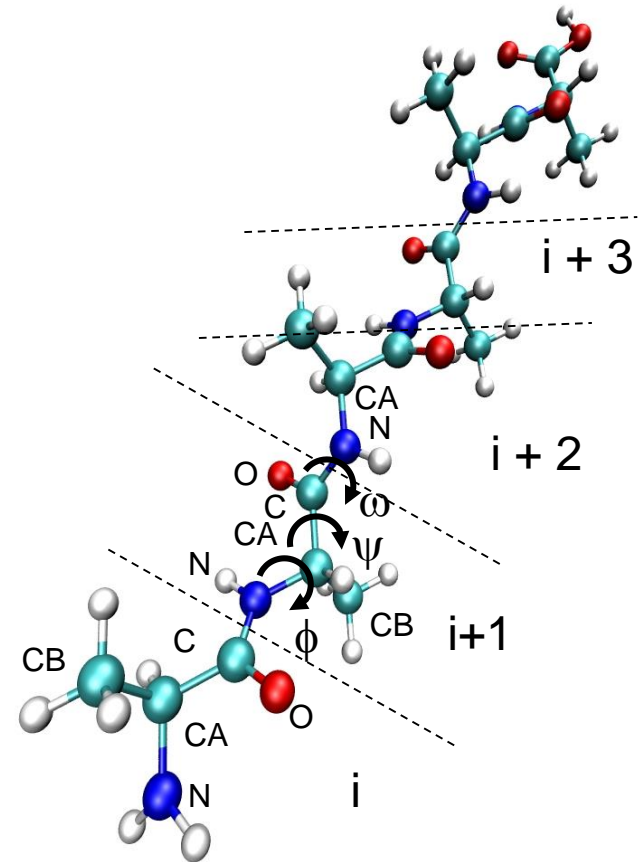


Na cadeia polipeptídica a definição do enovelamento da proteína depende dos ângulos de torção anterior ao carbono alfa, chamado de ângulo ϕ (phi em inglês) (ϕ), e do ângulo de torção após o carbono alfa, chamado ângulo ψ (ψ). A análise da rotação desses ângulos levou à identificação de regiões permitidas, onde não há choques entre os átomos, e regiões não-permitidas, onde há choques entre os átomos.

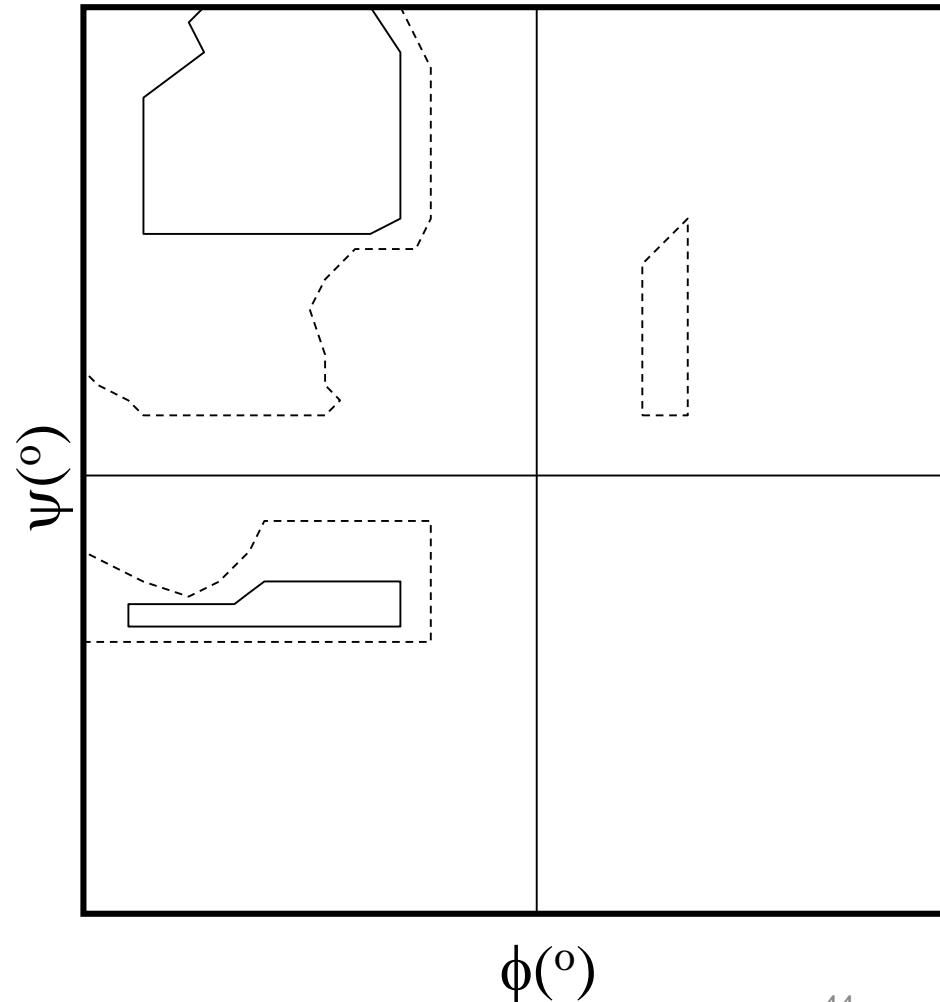


Gerando-se o gráfico para cada resíduo de aminoácido, o ângulo ϕ (fi) e ψ (psi), temos um diagrama bidimensional, onde as regiões permitidas e proibidas são identificáveis, tal diagrama é chamado de diagrama de Ramachandran. A cadeia principal apresenta um terceiro ângulo de torção, chamado ω (ômega). Este ângulo envolve a ligação parcialmente dupla entre o carbono da carbonila do resíduo de aminoácido $i+1$, com o nitrogênio do aminoácido $i+2$, ou seja, a ligação peptídica entre dois resíduos de aminoácido.

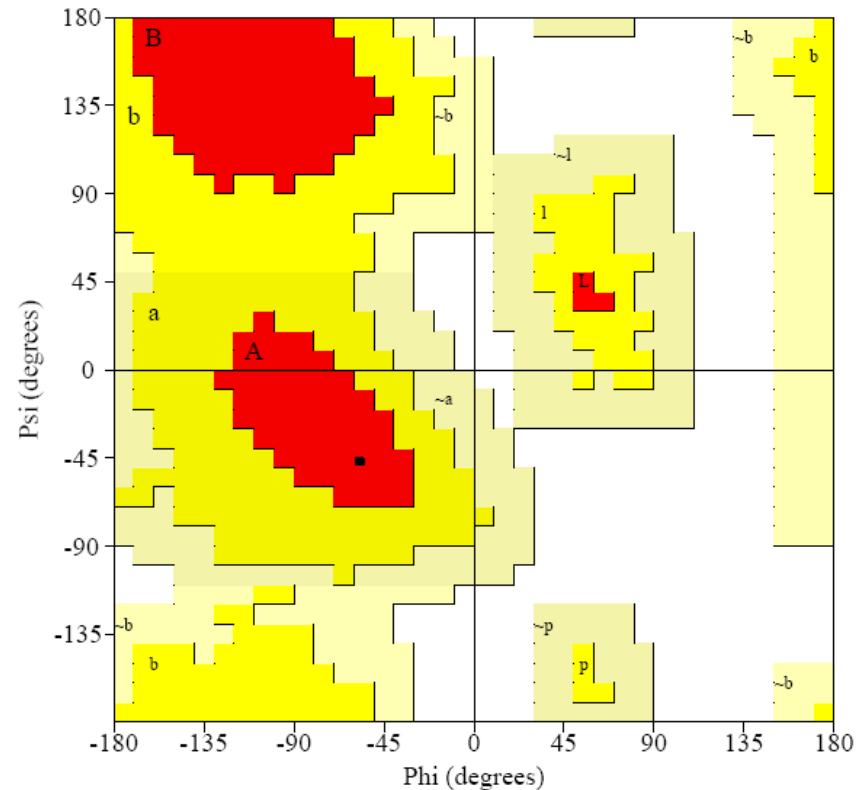
Devido ao caráter parcialmente duplo desta ligação, não há a liberdade estrutural observada para os ângulos ϕ (fi) e ψ (psi). O ângulo ω (ômega) admite duas situações, *trans*, onde seu valor é 180° e *cis*, onde seu valor é 0° .



Os eixos, ϕ (ϕ) e ψ (ψ), no diagrama de Ramachandran, variam de -180° a $+180^\circ$, as regiões no interior das áreas demarcadas no gráfico são regiões permitidas. Fica claro, a partir da análise do gráfico, que a área não permitida é maior do que a permitida. As regiões, não permitidas são possíveis de ocupação para a glicina, pois sua cadeia lateral restringe-se a um átomo de hidrogênio, permitindo mais liberdade para os ângulos ϕ (ϕ) e ψ (ψ).

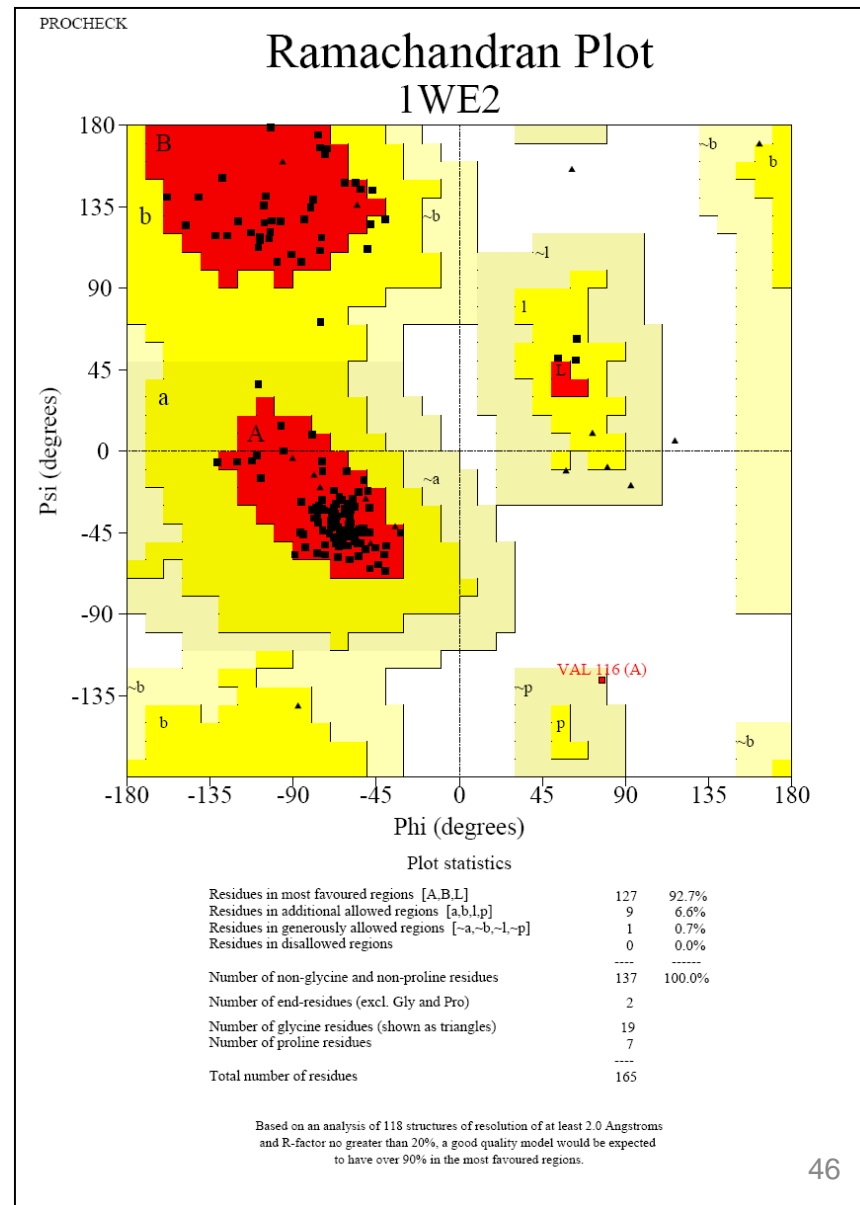


Em 1993 Laskowski e col. elaboraram o programa Procheck que define 4 regiões no diagrama de Ramachandran, são elas: região permitida (indicada em vermelho), região adicionalmente permitida (amarelo), região generosamente permitida (amarelo claro) e região proibida (branco). O programa Procheck usa dados da estrutura cristalográfica de 118 proteínas resolvidas a uma resolução melhor que 2,0 Å, para definir as regiões permitidas e proibidas do diagrama.

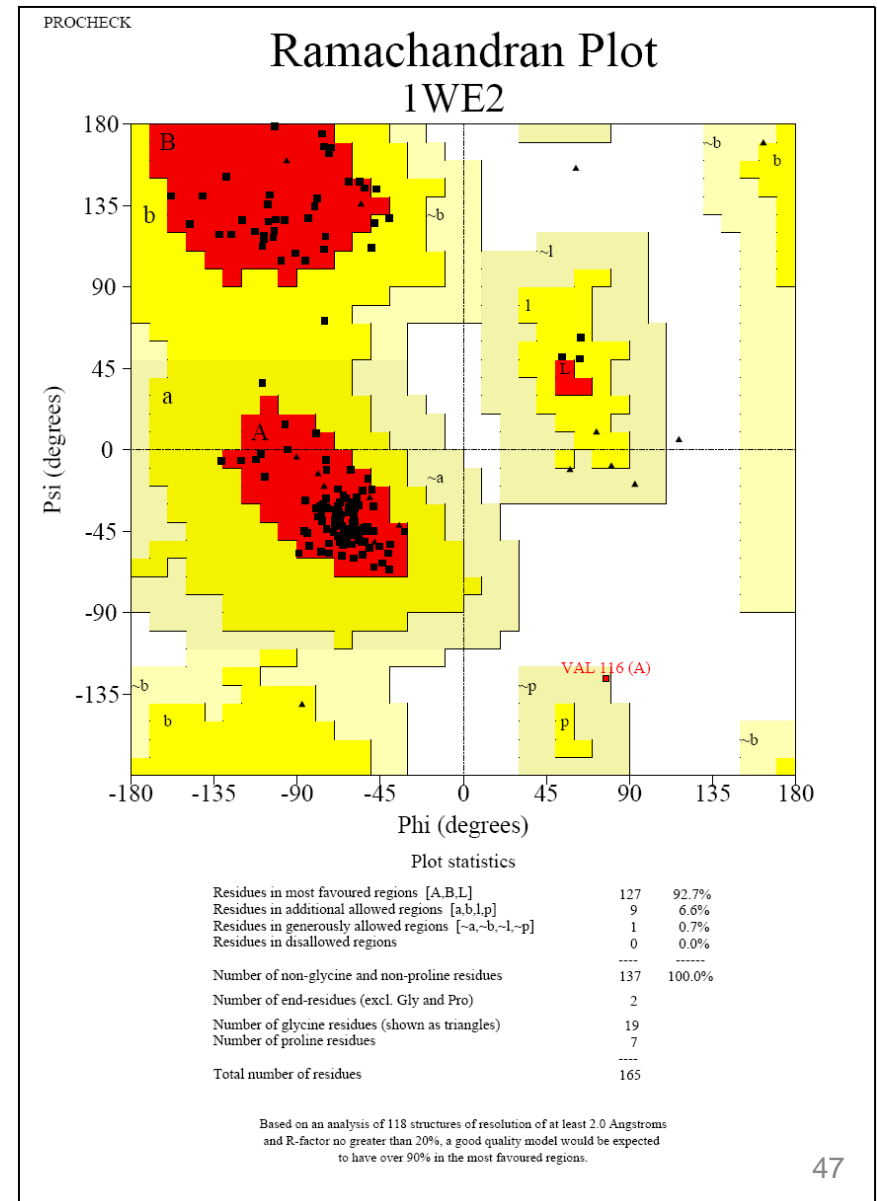


Referência: Laskowski, R.A.; MacArthur, M.W., Moss, D.S., Thornton, J.M. Procheck: a program to check the stereochemical quality of protein structures. *J. of Appl. Cryst.* 26(2), 283-291, (1993).

O programa Procheck indica as glicinas como triângulos, e estas podem ocupar qualquer região do diagrama de Ramachandran, visto que sua cadeia lateral restringe-se a um hidrogênio, que permite maior flexibilidade da cadeia principal. Os resíduos localizados nas regiões generosamente permitida e proibida são indicados em vermelho. Na figura ao lado o resíduo Val116 está na região generosamente permitida do gráfico.

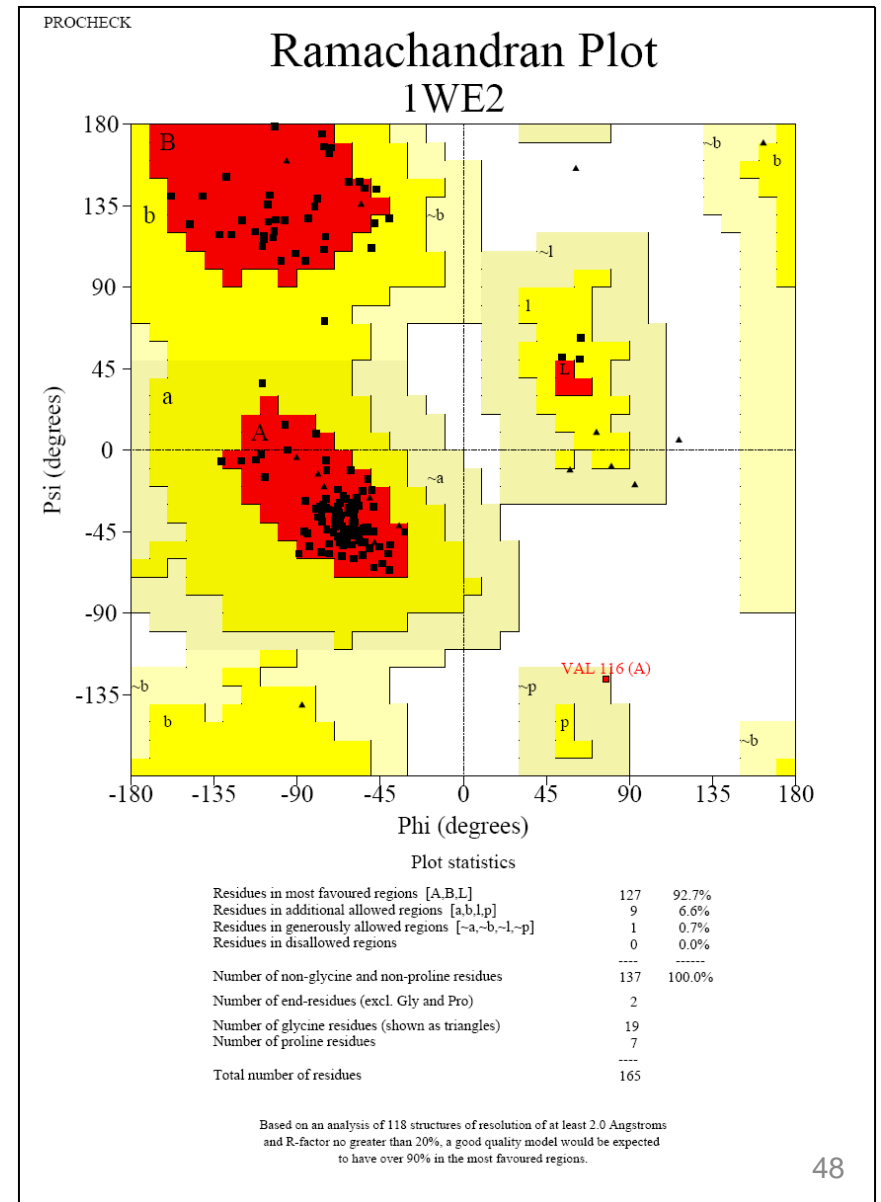


O Procheck oferece diversas informações, em outros gráficos, ou mesmo no gráfico principal. Uma delas é a estatística geral dos ângulos ϕ (fi) e ψ (psi), indicando a porcentagem de resíduos de aminoácido em cada região, na estrutura 1WE2 (Pereira *et al.*, 2004) temos 92,7 % dos resíduos da região permitida, 6,6 % na região adicionalmente permitida, e 0,7 % na região adicionalmente permitida. Para estruturas resolvidas a resolução acima de 2,0 Å espera-se acima de 90 % dos resíduos nas regiões permitidas.



Pereira, J.H., Oliveira, J. S., Canduri, F., Dias, M.V.B., Palma, M. S., Basso, L. A., Santos, D. S., & De Azevedo, W.F. Structure of shikimate kinase from *Mycobacterium tuberculosis* reveals the binding of shikimic acid. *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **60**, 2310-2319, 2004.

A análise da qualidade estereoquímica de modelos de proteínas é uma ferramenta poderosa na análise estrutural, sejam obtidos experimentalmente ou obtidos por modelagem molecular. O Programa Procheck facilita a análise dos modelos estruturais. Há versões do Procheck para Mac OS X, Windows e Linux, bem como, sites dedicados a análise on-line de proteínas.

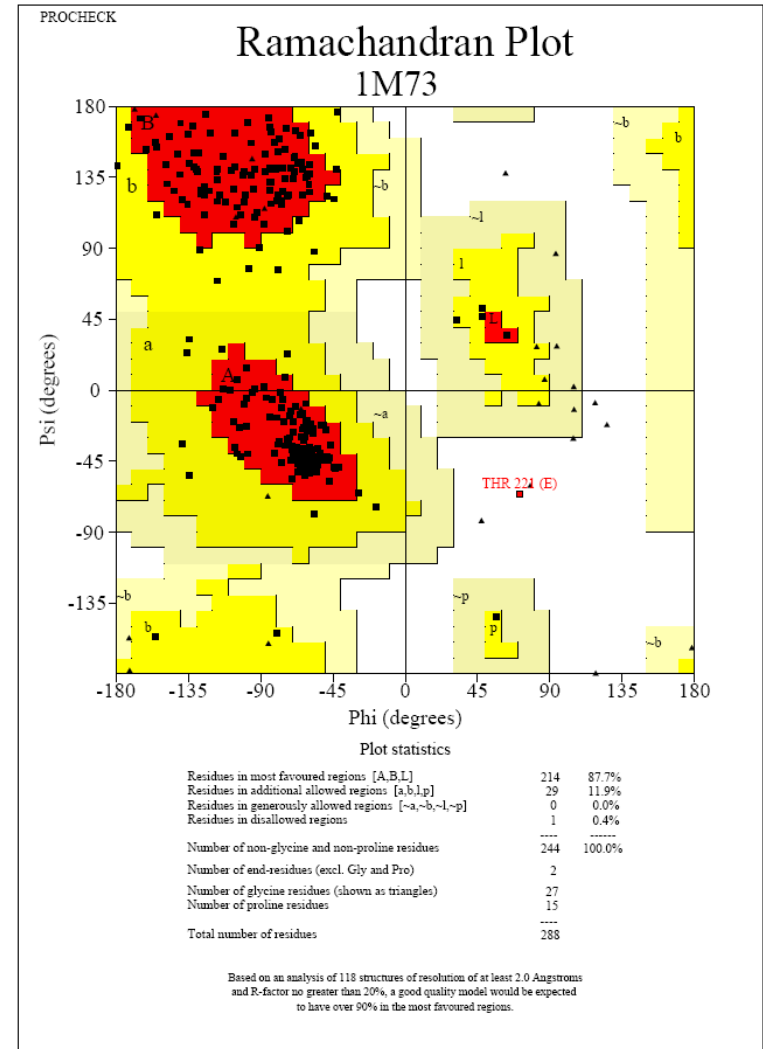
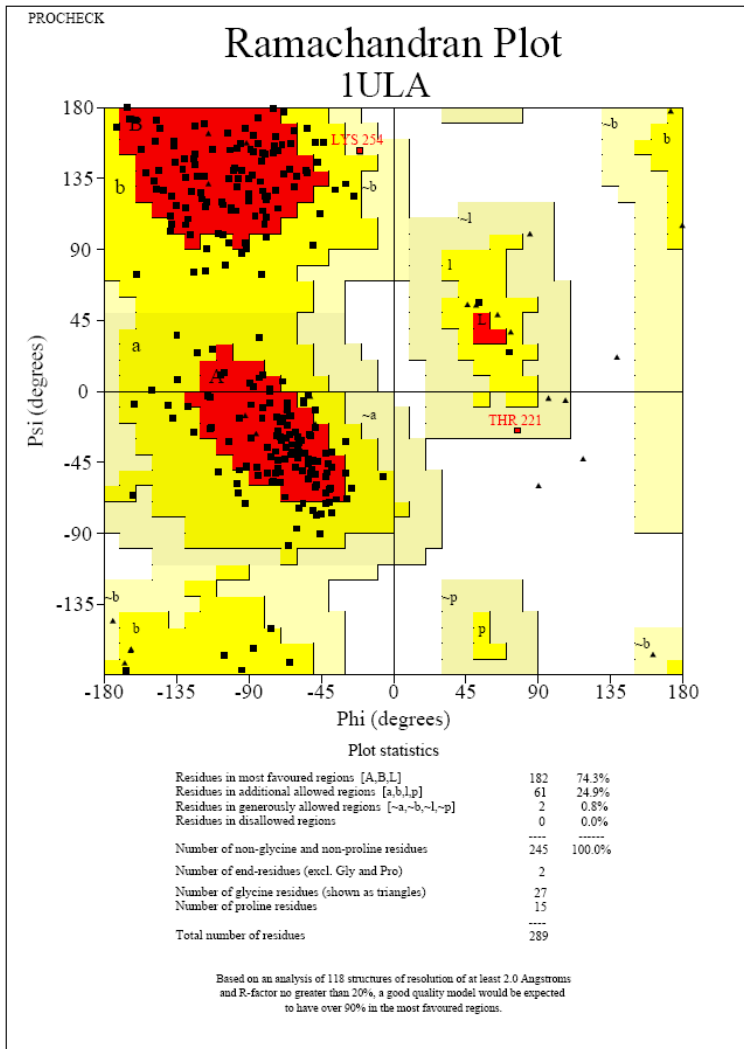


Como exemplo do uso do programa Procheck na análise da qualidade estereoquímica de modelos estruturais de proteínas, considere a proteína humana Purina Nucleosídeo Fosforilase (PNP, EC. 2.4.2.1). A PNP foi resolvida inicialmente em 1990 por Ealick *et al.* 1990 (código de acesso no PDB: 1ULA), a 2,75 Å de resolução. Uma nova estrutura foi refinada a 2,3 Å (De Azevedo *et al.*, 2003) (Código de acesso PDB: 1M73).

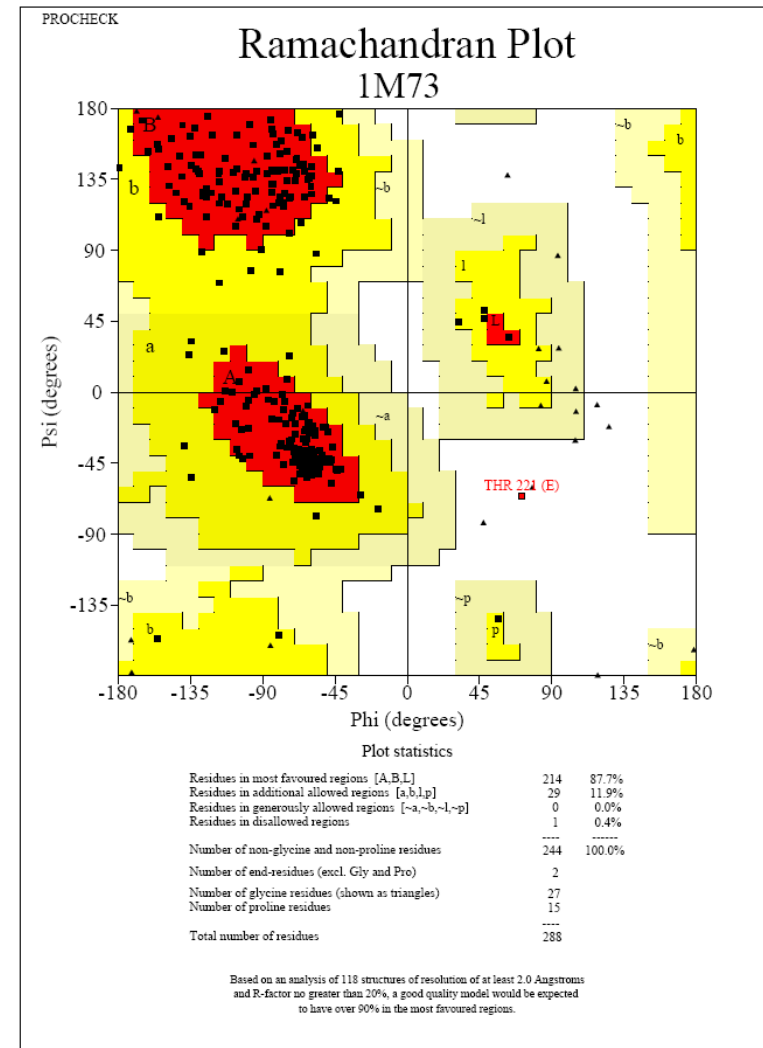
Referências:

De Azevedo, W. F., Canduri, F., Santos, D. M., Silva, R. G., Oliveira, J. S., Carvalho, L. P. S., Basso, L. A., Mendes, M. A., Palma, M. S., and Santos, D. S. Crystal structure of human purine nucleoside phosphorylase at 2.3 Å resolution. *Biochem. Biophys. Res. Commun.*, **308**(3), 545-552, 2003.

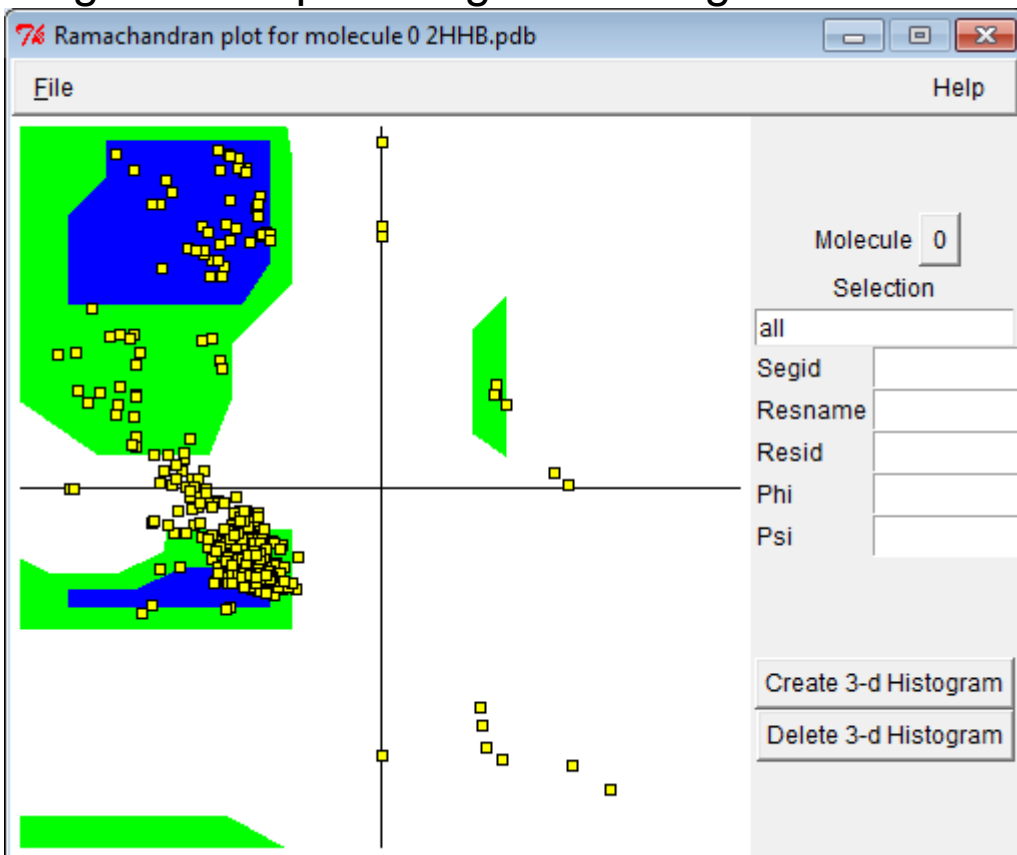
Ealick, S. E., Rule, S. A., Carter, D. C., Greenhough, T. J., Babu, Y. S., Cook, W. J., Habash, J., Helliwell, J. R., Stoeckler, J. D., Parks, R. E., Jr., Chen, S., -F., and Bugg, C. E. (1990). Three-dimensional structure of human erythrocytic purine nucleoside phosphorylase at 3.2Å resolution. *J. Biol. Chem.* 265(3), 1812-1820.



A análise dos dois diagramas de Ramachandran, indica claramente que a estrutura de coordenadas atômicas 1M73 apresenta melhor estatística estereoquímica. Na realidade, uma análise detalhada da estrutura tridimensional, resolvida de 2,3 Å, indicou que a estrutura 1ULA apresentava diversos erros, inclusive no sítio ativo. A estrutura 1ULA previa a participação de Lys244 no sítio ativo da enzima, a estrutura a mais alta resolução revelou que esta lisina estava a mais de 9 Å da posição inicialmente prevista. A análise do gráfico de Ramachandran, por si só, não valida a estrutura de uma proteína, mas é uma ferramenta valiosa na análise da estrutura 3D.



O programa VMD tem a opção de gerar o diagrama de Ramachandran para uma estrutura de proteína carregada no sistema. Após carregar um arquivo PDB com a estrutura de uma proteína clique nas seguintes opções no *VMD Main: Extensions>Analysis>Ramachandran Plot*. Você terá o gráfico de Ramachandran na tela, clique em *Molecule* para selecionar o arquivo PDB para o qual será gerado o diagrama. Depois de gerado o digrama teremos o gráfico, como mostrado abaixo.



Cada quadrado amarelo indica os ângulos phi e psi de cada aminoácido presente na estrutura. Se clicarmos sobre o quadrado amarelo teremos o valor dos ângulos de torção mostrados na tela.

Outro parâmetro, comumente utilizado para a análise da qualidade de um modelo estrutural, é o desvio médio quadrático da estrutura ideal, ou simplesmente RMSD (*root mean square deviation*). Este parâmetro mostra se as distâncias interatômicas (RMSD de ligação), os ângulos de ligação (RMSD de ângulo de ligação) e os ângulos de torção (RMSD de ângulo de torção) estão dentro de valores esperados. A partir de um conjunto de valores esperados ($d_{\text{referência}}$) para cada um dos parâmetros determinados experimentalmente ($d_{\text{experimental}}$) o cálculo do RMSD é efetuado. Os valores de referência são tirados de estruturas de pequenas moléculas, que são resolvidas normalmente a mais alta resolução que as estruturas de proteínas. A somatória é feita para o número de ocorrência do parâmetro. Por exemplo, a distância do N ao CA, tem um valor esperado ($d_{\text{referência}}$) e a soma é feita para todas as distâncias N-CA encontradas na estrutura.

$$RMSD = \sqrt{\frac{1}{N} \sum_{i=1}^N (d_{\text{experimental}} - d_{\text{referência}})^2}$$

DRENTH, J. (1994). *Principles of Protein X-ray Crystallography*. New York: Springer-Verlag.

LESK, A. M. Introduction to Protein Architecture. Oxford University Press, New York, 2001.

RHODES, G. (2000). *Crystallography Made Crystal Clear*. 2nd ed. San Diego: Academic Press.

STOUT, G. H. & JENSEN, L. H. (1989). *X-Ray Structure Determination. A Practical Guide*. 2nd ed. New York: John Wiley & Sons.