

Curso de Modelagem Molecular

Módulo II: Métodos de Cálculo e Aplicações

Dr. Carlos Mauricio R. Sant'Anna

GPAQ, Dequim, UFRuralRJ

LASSBio, FF, UFRJ

the past decade as substance P antagonists, a large percentage of which likely exhibited sub-nanomolar potency at the NK-1 receptor, but only one has made it to market. This example highlights the difficulty of and resources needed to optimize the ancillary properties of potent inhibitors/antagonists so that they can become safe, viable medicines.

In the drug discovery process, we must also be cognizant of the interrelatedness of academic, government, and industrial research in the development of new drug entities. Despite large research budgets, the biomedical research carried out by pharmaceutical companies still represents only a small percentage of the overall worldwide research effort on diseases and approaches to their treatment. Academic and government laboratories, funded with public monies, often provide much basic research and fundamental insight into diseases that can direct researchers toward novel ways of attacking diseases. However, they are rarely organized (nor is it their mission) to embrace the drug discovery process in the multidisciplinary fashion outlined above that is the modern paradigm by which new hits or leads are first identified and then get transformed into new viable medicines. All of the above discussion speaks to one of the most

important issues facing discovery medicinal chemistry today: the continuing need for excellent synthetic chemists. In large pharmaceutical companies, the drug discovery process is driven by multidisciplinary teams made up of the very best experts in each discipline, and chemistry is one key element in this. These teams have ready access to experts in other areas of biomedical science, and although chemists often end up as group leaders of discovery efforts, that usually occurs after much experience has been gained in the drug discovery process. The recent advances discussed above have put more tools in the chemist's toolkit, but in order to use these tools effectively, it inevitably comes down to the ability to make the absolutely "correct" molecule in a timely and cost-effective manner. This process requires the very best organic chemistry skills, and we must continue to provide funding in the university system for training in these core skill sets to chemists in their graduate and postdoctoral studies if we are to continue to provide the very best in medicines for what is becoming an aging population.

References and Notes

1. In this discussion, a "hit" is defined as a nonoptimized structure obtained from some screening process on a target protein. It is often a very weak binder and is likely to have a nonoptimized pharmacokinetic profile. A

- "lead" is defined as a structure that has been derived from an early "hit" and, although still not fully optimized, has been shown to have some appropriate characteristics to be a precursor of a drug entity. Often a good lead will have shown some proof-of-concept activity in an *in vivo* pharmacological model, but will likely not have been fully optimized for pharmacokinetic properties or undesirable off-target activities.
2. C. A. Lipinski, F. Lombardo, B. W. Dominy, P. J. Feeney, *Adv. Drug Delivery Rev.* 23, 3 (1997).
3. In the discovery setting, the rule of five (2) predicts that poor absorption or permeation of drugs is more likely when a drug molecule possesses either (i) more than 5 hydrogen bond donors, (ii) 10 hydrogen bond acceptors, (iii) a molecular weight greater than 500, or (iv) a calculated $\log P$ greater than 5.
4. J. Uetrich, *Drug Discov. Today* 8, 932 (2003).
5. D. C. Evans, A. R. Watt, D. A. Nicol-Grimith, T. A. Ballie, *Chem. Rev.* 104, 17, 3 (2004).
6. See (17) for an excellent review of the cardiovascular effects manifested by QT interval prolongation and the evaluation of drug candidates for this parameter.
7. See the cover story in *Drug Discov. Dev.* 6, 30 (2003).
8. T. Koppal, *Drug Discov. Dev.* 6, 50 (2003).
9. A. DePalma, *Drug Discov. Dev.* 5, 50 (2002).
10. A. DePalma, *Drug Discov. Dev.* 6, 51 (2003).
11. H. McCoy, J. F. Tremblay, *Chem. Eng. News* 81, 15 (2003).
12. A. M. Rouh, *Chem. Eng. News* 81, 75 (2003).
13. T. Koppal, *Drug Discov. Dev.* 6, 22 (2003).
14. W. A. Korfmecher et al., *Rapid Commun. Mass Spectrom.* 15, 335 (2001).
15. J. J. Hale et al., *J. Med. Chem.* 41, 4607 (1998).
16. R. M. Smider et al., *Science* 251, 435 (1991).
17. R. Metzler, A. Ebneth, U. Borchert, Q. Pongs, *Drug Discov. Today* 6, 78 (2001).

REVIEW

The Many Roles of Computation in Drug Discovery

William L. Jorgensen

An overview is given on the diverse uses of computational chemistry in drug discovery. Particular emphasis is placed on virtual screening, *de novo* design, evaluation of drug-likeness, and advanced methods for determining protein-ligand binding.

"Is there really a case where a drug that's on the market was designed by a computer?" When asked this, I invoke the professorial mantra ("All questions are good questions."), while sensing that the desired answer is "no". Then, the inquisitor could go back to the lab with the reassurance that his or her choice to avoid learning about computational chemistry remains wise. The reality is that the use of computers and computational methods permeates all aspects of drug discovery today. Those who are most proficient with the computational tools have the advantage for delivering new drug candidates more quickly and at lower cost than their competitors.

However, the phrasing of the question

suggests misunderstanding and oversimplification of the drug discovery process. First, it is the rare case today when an unmodified natural product like taxol becomes a drug. It is also inconceivable that a human with or without computational tools could propose a single chemical structure that ends up as a drug; there are far too many hurdles and subtleties along the way. Most drugs now arise through discovery programs that begin with identification of a biomolecular target of potential therapeutic value through biological studies including, for example, analysis of mice with gene knockouts. A multidisciplinary project team is then assembled with the goal of finding clinical candidates, i.e., druglike compounds that are ready for human clinical trials, which typically selectively bind to the molecular target and interfere either with its activity as a recep-

tor or enzyme. Molecular libraries are screened, and the resulting leads are optimized in a cycle that features design, synthesis and assaying of numerous analogs, and animal studies. Crystal structure determination for complexes of some analogs with the biomolecular target is often possible, which enables "structure-based drug design" (SBDD) and the efficient optimization of leads. The success of SBDD is well documented (1, 2); it has contributed to the introduction of ~50 compounds into clinical trials and to numerous drug approvals. Minimally, the role of computation here is in the structure refinement using simulated annealing (3), development of the underlying molecular mechanics (MM) force fields, structure display, and building and MM evaluation of analogs. All top pharmaceutical companies have substantial structural biology and computational chemistry groups that are intertwined and participate on the project teams.

There is usually much "tweaking" toward the end of the preclinical period of drug discovery when a series of compounds

"Is there really a case where a drug that's on the market was designed by a computer? When asked this, I invoke the professorial mantra ("All questions are good questions."), while sensing that the desired answer is "no". Then, the inquisitor could go back to the lab with the reassurance that his or her choice to avoid learning about computational chemistry remains wise."

"...the phrasing of the question suggests misunderstanding and oversimplification of the drug discovery process."

"There is not going to be a "voila" moment at the computer terminal. Instead, there is systematic use of wide ranging computational tools to facilitate and enhance the drug discovery process."

www.sciencemag.org SCIENCE 303
19 MARCH 2004

Department of Chemistry, Yale University, New Haven, CT 06520-8107, USA. E-mail: william.jorgensen@yale.edu

Estratégias Gerais para Obtenção de Compostos Bioativos

"Screening" cego: avaliação da atividade de produtos naturais (plantas, animais marinhos, processos fermentativos) ou compostos sintéticos (bibliotecas de estruturas ou novas estruturas obtidas com o uso de química combinatória)



"Lead"



Otimização estrutural

Planejamento racional: planejamento a partir do estudo direto ou indireto do mecanismo de interação entre o composto bioativo (ligante) e o alvo biológico

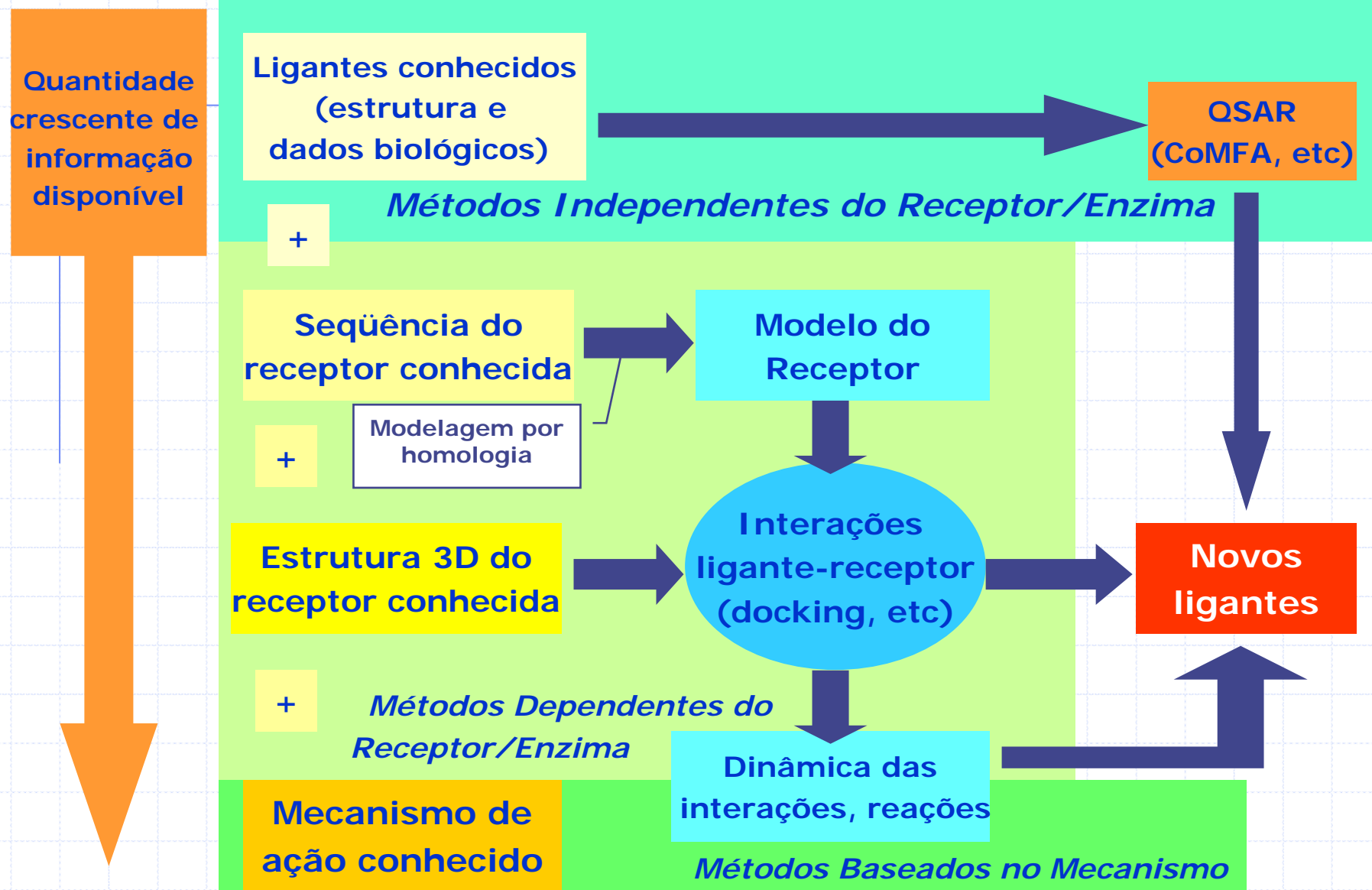


"Lead"

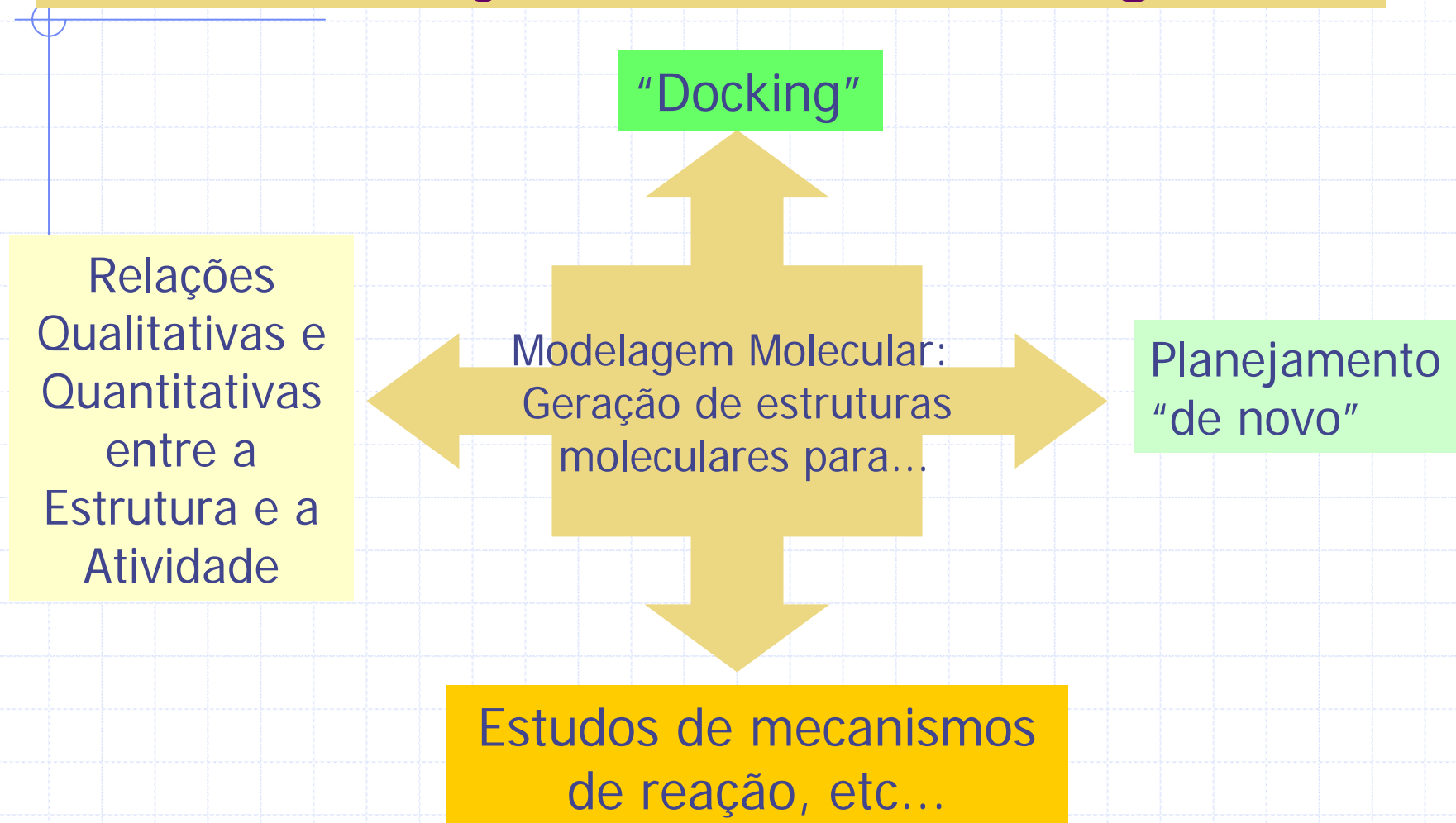


Otimização estrutural

Estratégias no Planejamento Racional



Modelagem Molecular no Planejamento de Bioligantes



Modelagem Molecular: Classificações dos Métodos

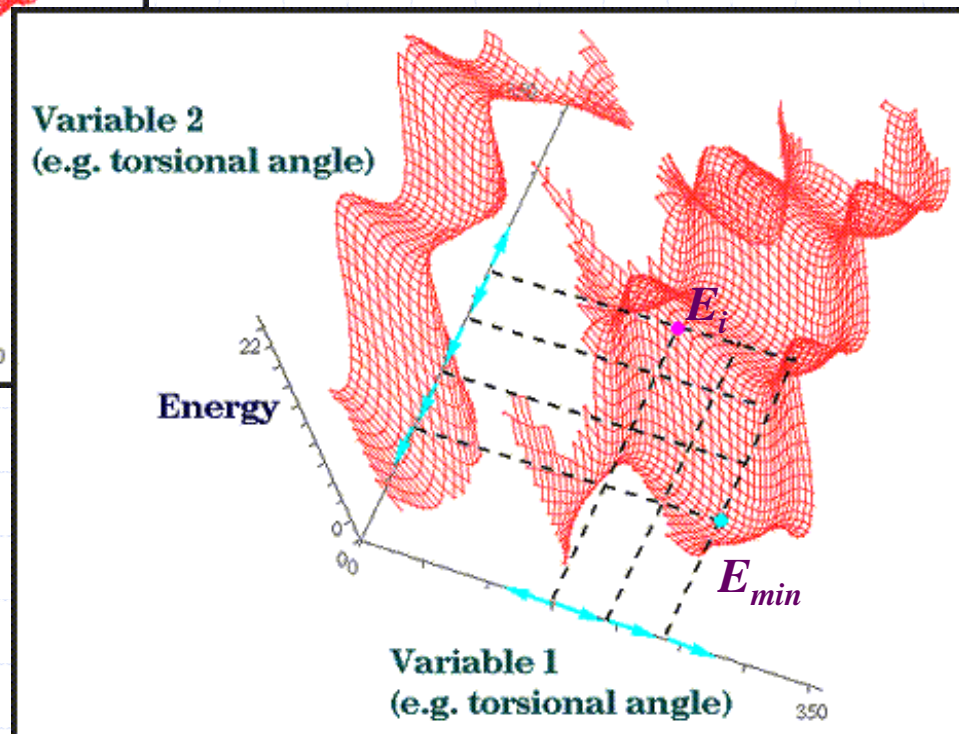
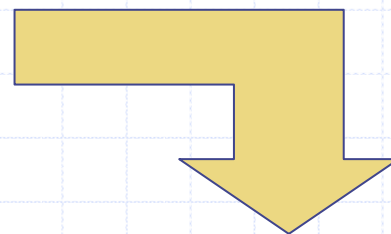
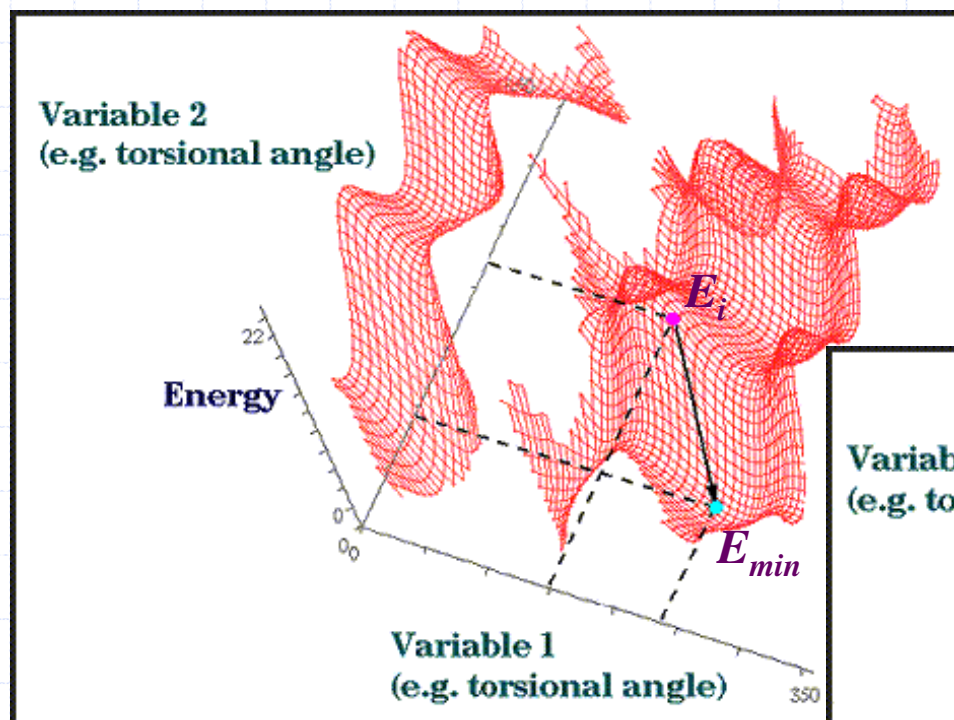
Quanto à Dinâmica do Método de Cálculo da Energia

- Métodos de otimização estrutural direta (busca do ponto estacionário mais próximo)
- Métodos de otimização múltipla (métodos estatísticos, análise sistemática...)
- Métodos dinâmicos: geração de trajetórias (dinâmica molecular...)

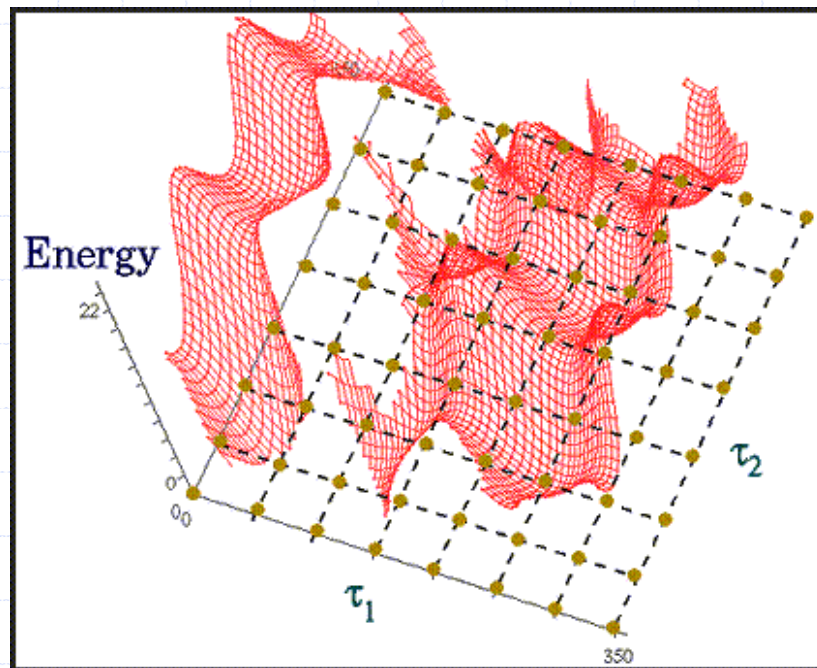
Quanto ao Tipo do Método de Cálculo da Energia

- Métodos de Campo de Força: mecânica molecular
- Métodos Quânticos: Semi-empíricos, *ab initio* HF, MP, DFT
- Métodos de Simulação: métodos de energia livre, métodos de perturbação termodinâmica, métodos de integração termodinâmica

Busca Direta de Mínimos de Energia

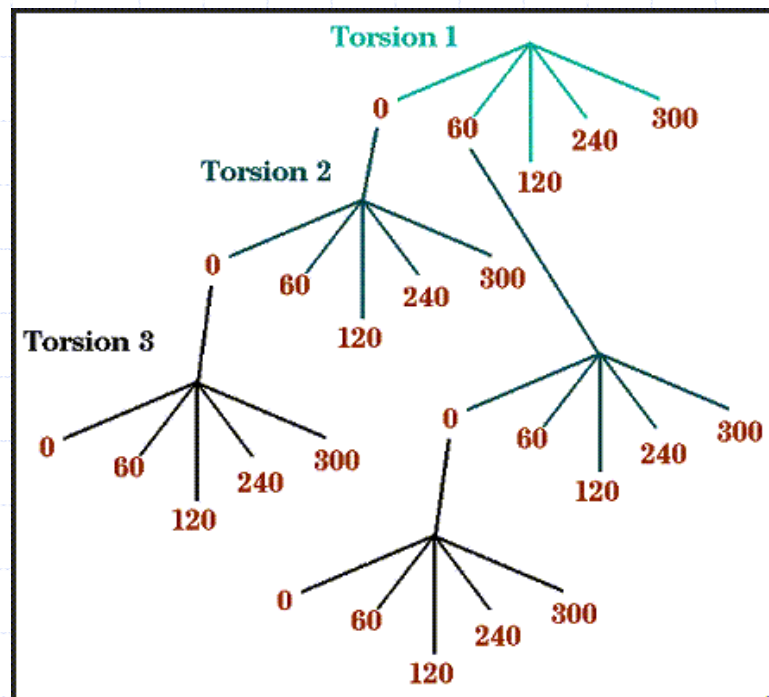


Análise Conformacional Sistemática: Pesquisa de Grade

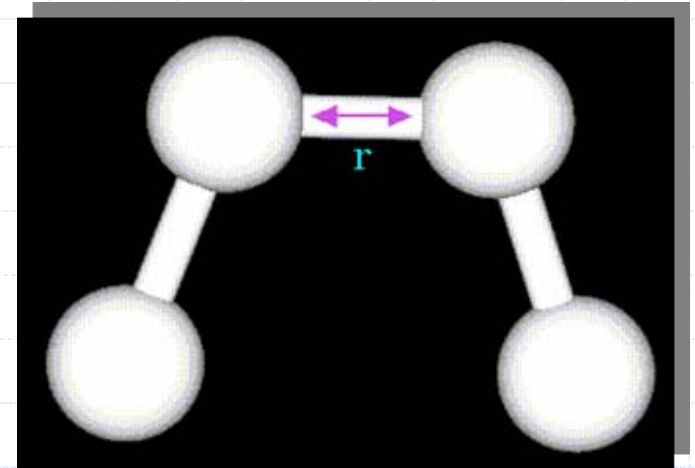
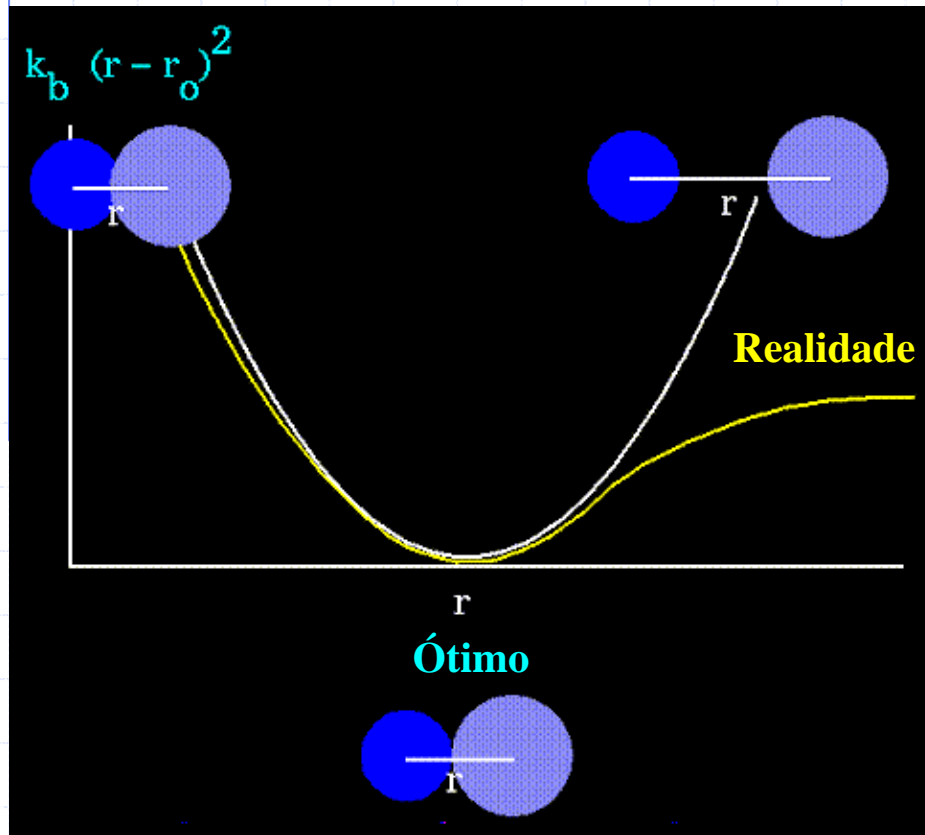


O número de confômeros equivale a $(360^\circ/\theta)^n$, onde θ é o incremento usado no processo de varredura e n é o número de ligações avaliadas.

O número de conformações necessárias para varrer adequadamente todo o espaço conformacional de uma molécula já se torna muito grande quando o número de ligações com rotação livre **excede 5**.

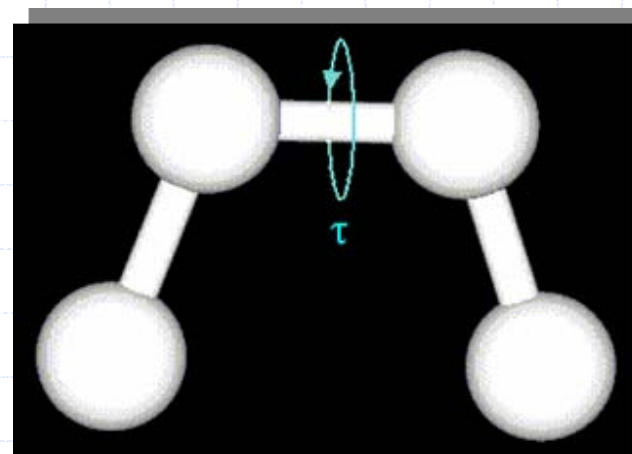
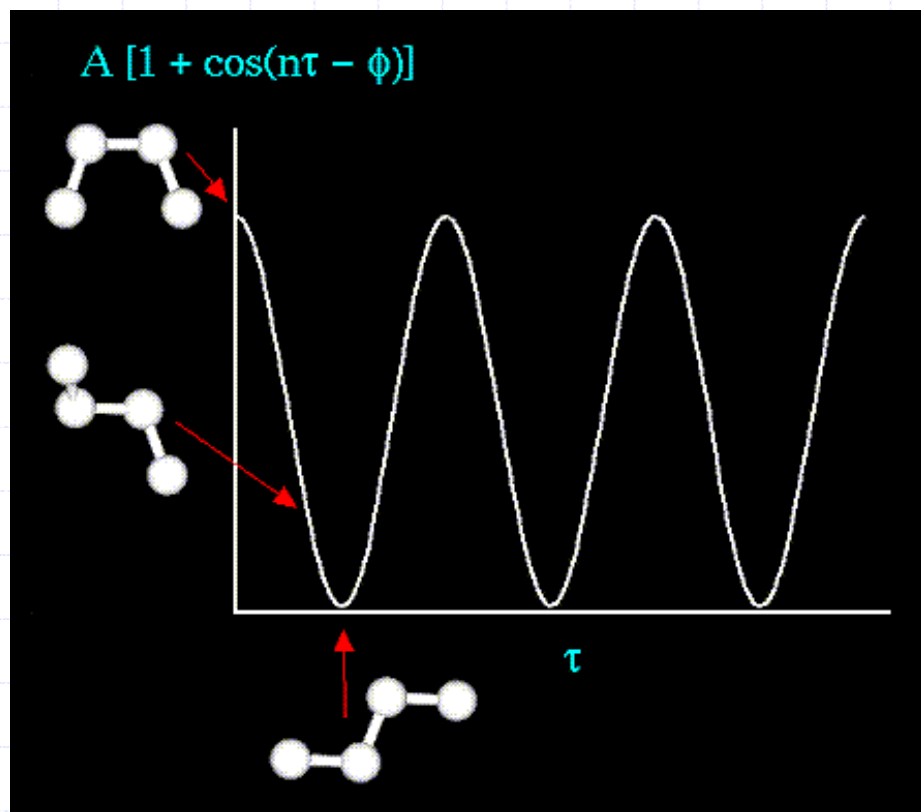


Estrutura e Energia: Métodos Empíricos



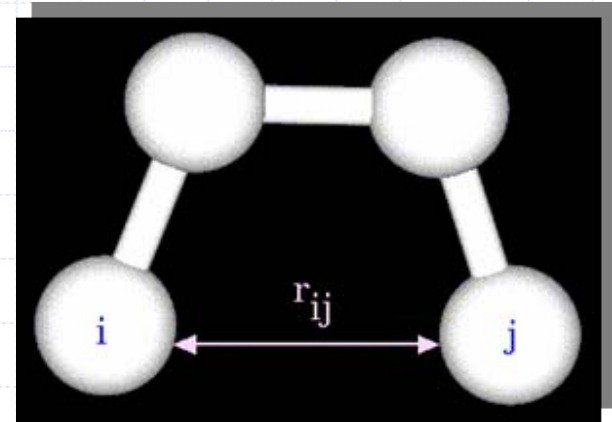
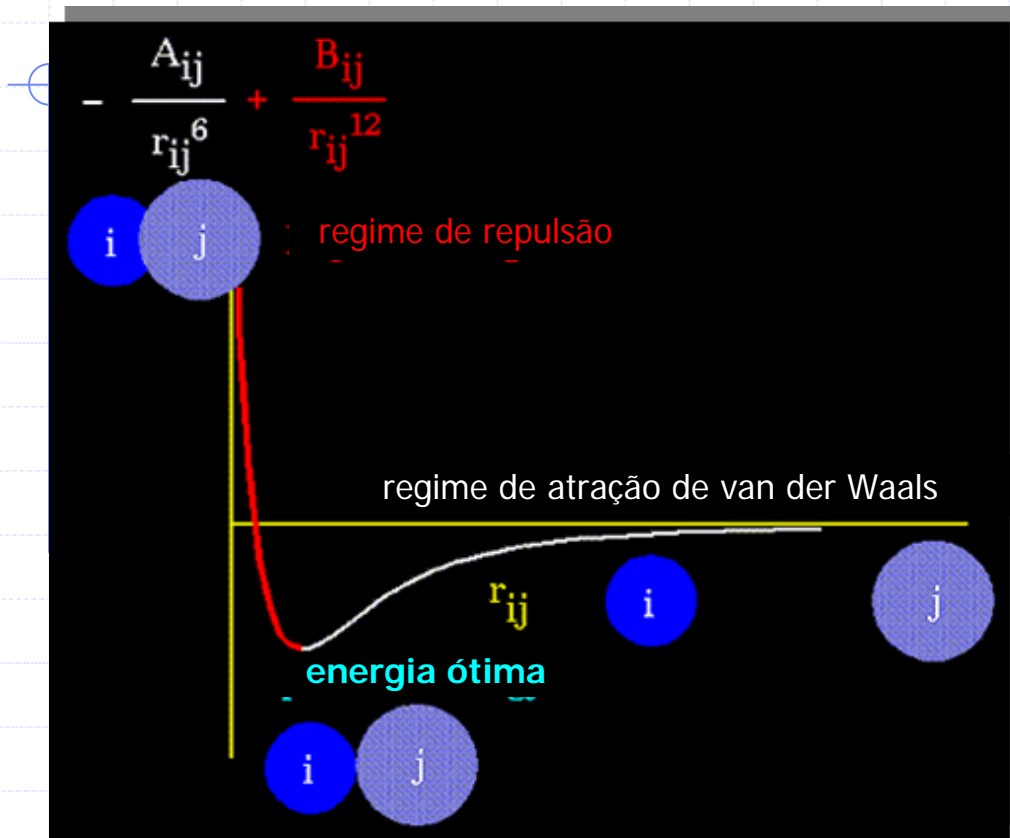
$$E = \sum_{\text{ligações}} k_b (r - r_o)^2$$

Estrutura e Energia: Métodos Empíricos



$$E = \sum_{\text{diedros}} A [1 + \cos(n\tau - \phi)]$$

Estrutura e Energia: Métodos Empíricos



$$E = \sum_i \sum_j \left(\frac{-A_{ij}}{r_{ij}^6} + \frac{B_{ij}}{r_{ij}^{12}} \right) + \sum_i \sum_j \frac{q_i q_j}{r_{ij}}$$

termo de van der Waals termo eletrostático

Métodos Empíricos: Campos de Força

Campo de Força	Sistemas Moleculares
EFF	Alcanos
MM3	Geral
CVFF	Geral
Tripes	Geral
MMFF	Geral
Dreiding	Geral
Amber	Proteínas, ácidos nucleicos, carboidratos
CHARMM	Proteínas
GROMOS	Proteínas, ácidos nucleicos, carboidratos
MOMEK	Compostos de coordenação

Dinâmica Molecular: Energia Potencial e Energia Cinética

Termodinâmica \Rightarrow Quais estados são possíveis?

Cinética \Rightarrow Como (e com que velocidade) os estados se interconvertem?

Termodinâmica:	estado 1		estado 2		estado n
Cinética:	estado 1	\rightleftharpoons	estado 2	\rightleftharpoons	estado n

$$-\frac{dE}{dr_i} = F_i$$

$$-\frac{dE}{dr_i} = m_i \frac{d^2 r_i}{dt^2}$$

1 Resolver para a_i em t usando:

$$-\frac{dE}{dr_i} = F_i = m_i a_i(t)$$

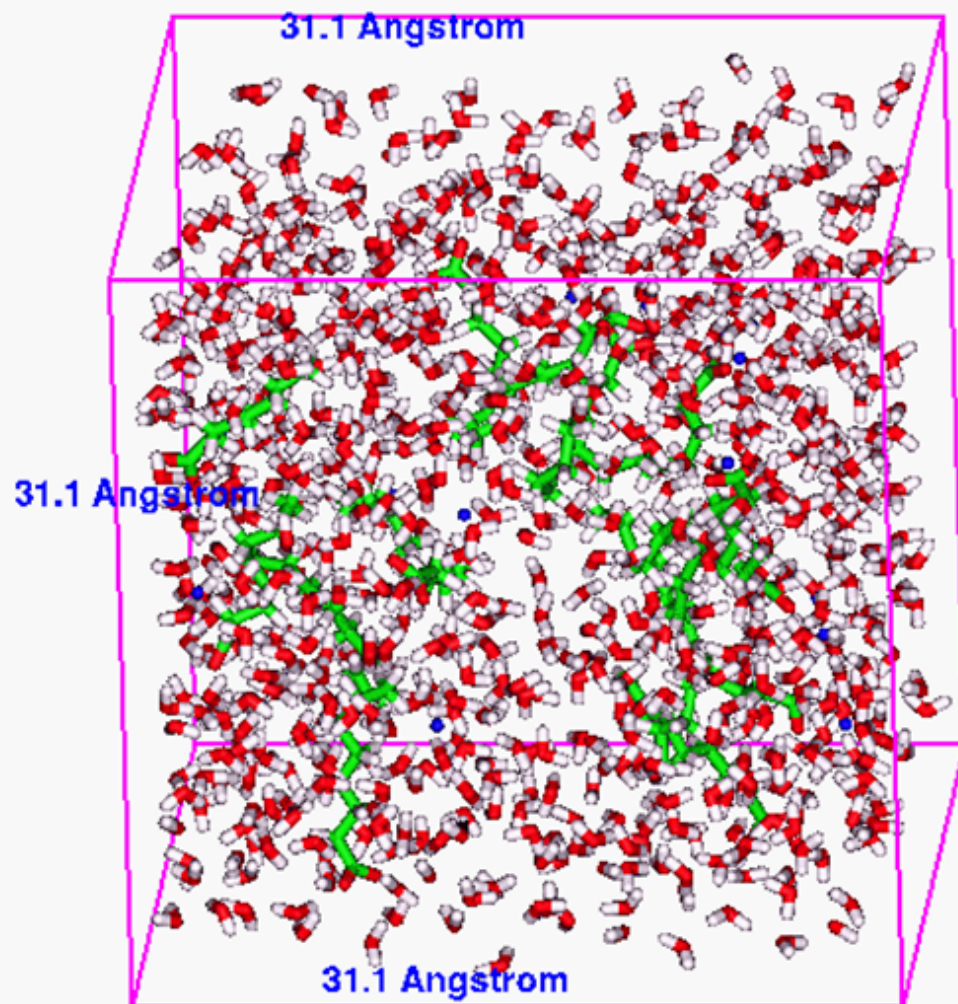
2 Atualizar v_i em $t + \Delta t/2$ usando:

$$v_i(t + \Delta t/2) = v_i(t - \Delta t/2) + a_i(t) \Delta t$$

3 Atualizar r_i em $t + \Delta t$ usando:

$$r_i(t + \Delta t) = r_i(t) + v_i(t + \Delta t/2) \Delta t$$

Dinâmica Molecular



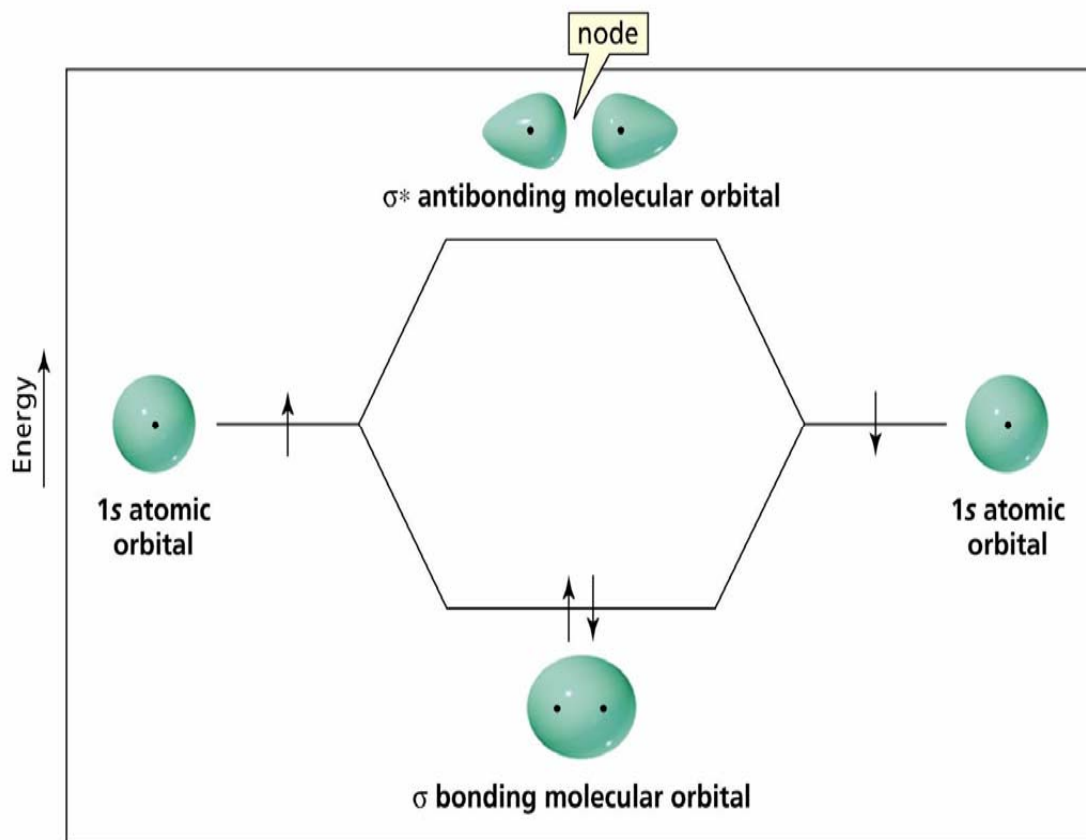
Estudo por
dinâmica
molecular da
formação de
micelas em água:
15 moléculas de
octanoato de
sódio + 792
moléculas de
água

Estrutura e Energia: Métodos Quânticos

• Nos métodos quânticos, a energia da estrutura molecular é definida colocando-se os núcleos em regiões nas quais os elétrons estão associados a orbitais, regiões do espaço definidas pela resolução da equação de Schrödinger:

$$E\Psi = H\Psi$$

onde Ψ é a chamada **função de onda** e E é a energia eletrônica.



*Representação de orbitais moleculares
para a molécula de H_2*

Métodos Quânticos: Diferenças

- A descrição de moléculas é feita pela resolução da equação de Schrödinger na qual a função de onda Ψ é uma função das coordenadas dos núcleos (R) e dos elétrons (r):

$$H\Psi(R,r) = E\Psi(R,r) \quad (1.1)$$

Como os núcleos são milhares de vezes mais pesados do que os elétrons, os movimentos nuclear e eletrônico podem ser considerados separadamente:

$$H\Phi(R) = E\Phi(R) \quad (1.2)$$

$$H\psi(r;R) = E\psi(r;R) \quad (1.3)$$

- A resolução de (1.2) é chamada *dinâmica quântica* e depende de uma função de energia (a SEP) definida apenas pelas posições dos núcleos, $E(R)$.

Métodos *ab initio* resolvem (1.3) diretamente.

Ex.: HF, DFT, MPx

Métodos *semi-empíricos* fazem aproximações que são compensadas com funções ajustadas empiricamente.

Métodos Quânticos: Diferenças

Métodos ab initio

1. Conjuntos de Bases

Mínimo: STO-3G

De valência dividida: 3-21G, 6-31G, 6-311G, ...

Com funções de polarização: 6-31G*, 6-311G**, ...

Com funções difusas: 6-31+G*, 6-311++G**, ...

2. Programas

Gaussian, GAMESS, Spartan

Métodos semi-empíricos

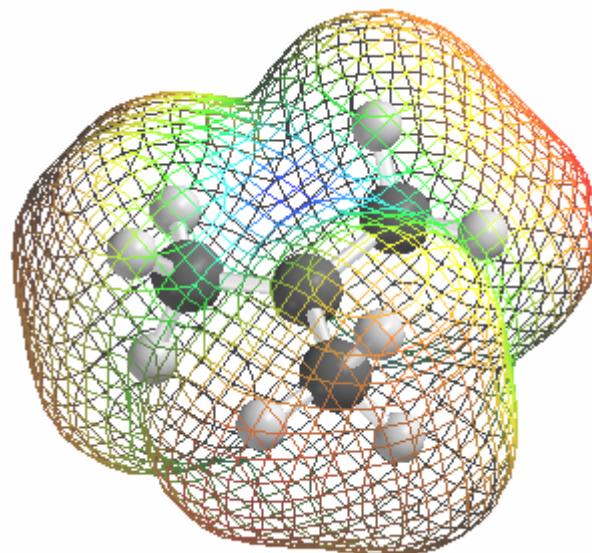
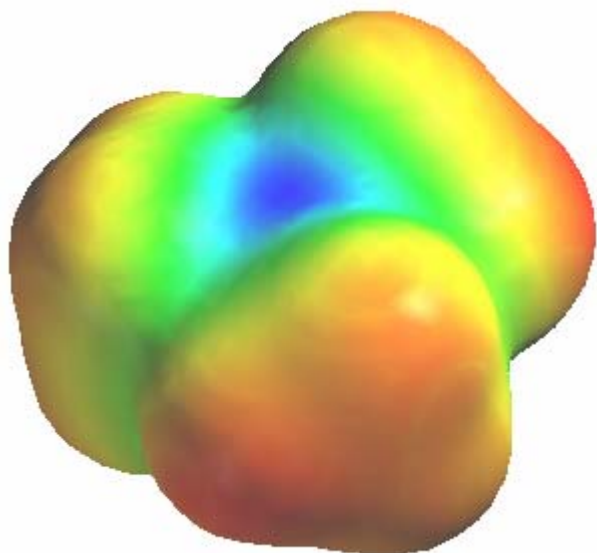
1. Hamiltonianos

PM6, RM1, AM1, PM3, MNDO, ...

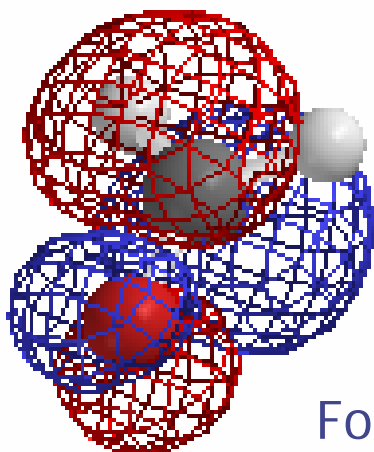
2. Programas

Mopac 6.0, Mopac 2003, Mopac 2007, Ampac, Spartan, Sybyl

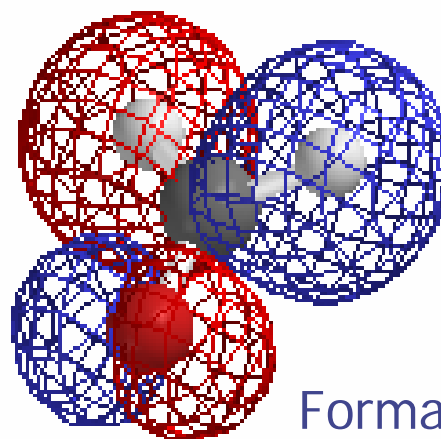
Aplicações: Propriedades Eletrônicas



Cátion *t*-butil:
mapas de
potencial
eletrostático

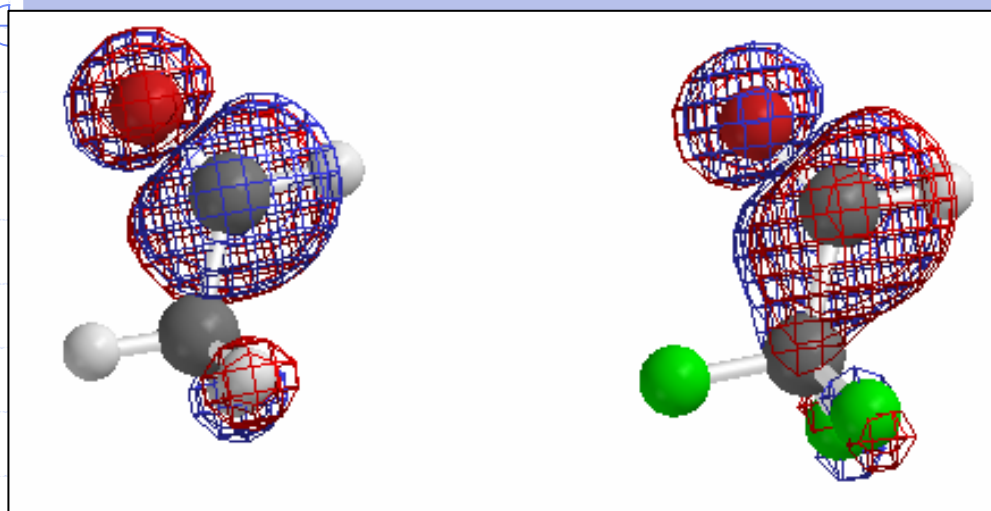


Formaldeído: LUMO

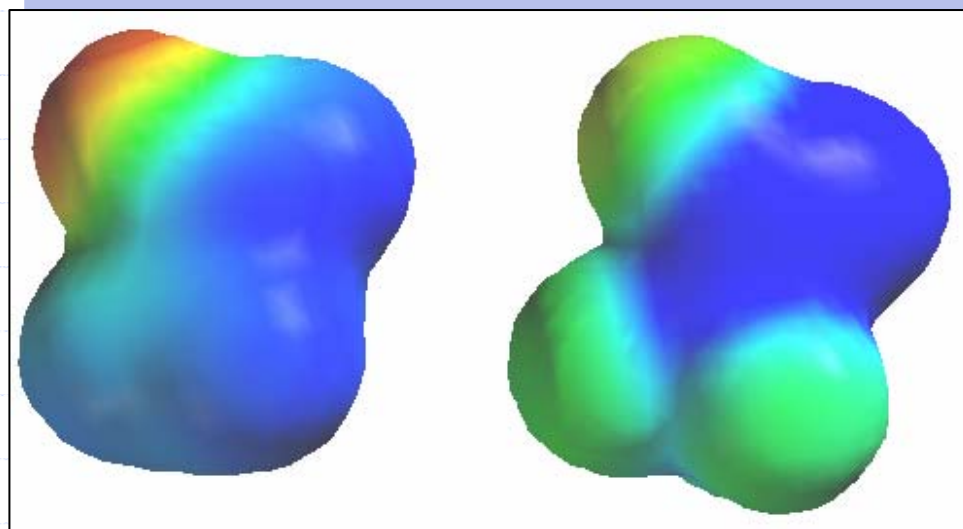


Formaldeído: HOMO

Aplicações: Estudo da Reatividade Química

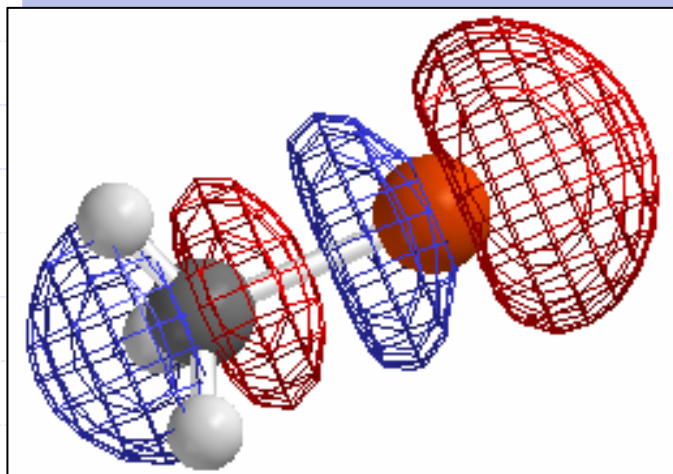


Representação do LUMO dos substratos acetaldeído e trifluór-acetaldeído (método quântico semi-empírico AM1)

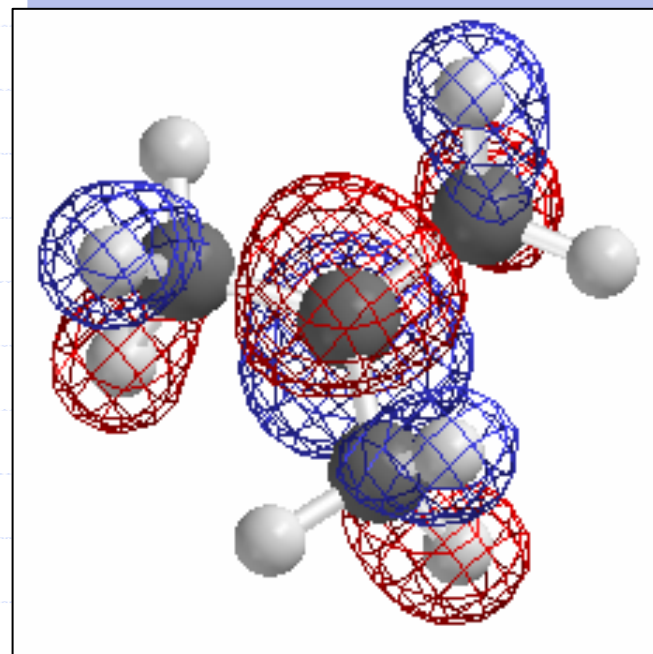


Representação superfície de potencial eletrostático dos substratos acetaldeído e trifluór-acetaldeído (método quântico semi-empírico AM1)

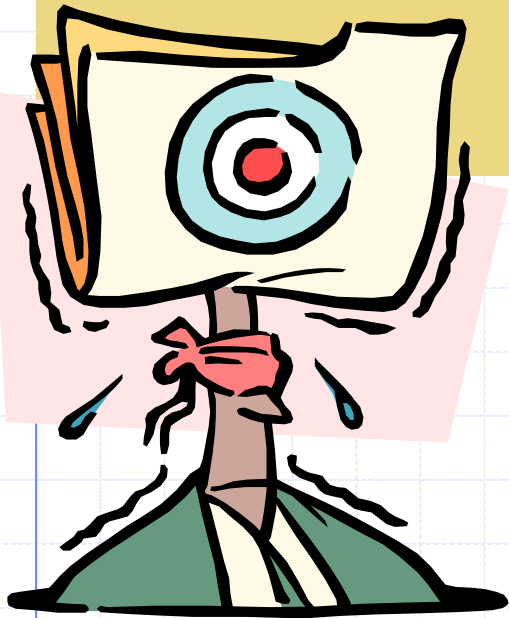
Aplicações: Estudo da Reatividade Química



S_N2 : representação do LUMO do substrato CH_3Br obtido pelo método quântico semi-empírico AM1



S_N1 : representação do LUMO do intermediário cátion *t*-butil obtido pelo método quântico semi-empírico AM1



Qual o Melhor Modelo?

Os principais parâmetros para a escolha de um modelo teórico são o **desempenho** e o **custo**.


O sucesso de qualquer modelo depende primeiro da sua capacidade de **reproduzir dados experimentais**.

Mas um modelo precisa também ser **prático**, o que depende do tamanho do sistema em estudo e das facilidades computacionais disponíveis.

Um **modelo prático** provavelmente não vai ser o melhor tratamento possível para um determinado problema, mas **aquele que produzirá as informações necessárias com uma precisão e um custo razoáveis**.

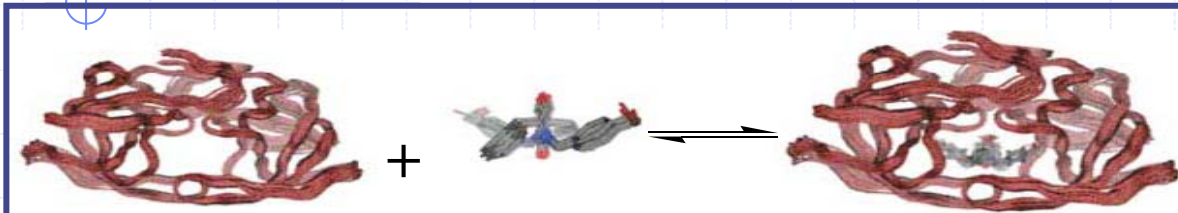
Desempenho Relativo dos Métodos de Cálculo de Modelagem Molecular

Objetivo	Mecânica Molecular	Semi-empírico	<i>Ab Initio</i>	
			HF	MP2
Geometria (grupo principal)	aceit.→bom	Bom	Bom	Bom
Geometria (metais de transição)	-	Bom	Pobre	?
Geometria (ests. de transição)	-	Bom	Bom	Bom
Conformação	Bom	Pobre	aceit.→bom	Bom
Termoquímica (não isodêmica)	-	Pobre	aceit.→bom	Bom
Termoquímica (isodêmica)	-	aceitável	Bom	Bom
Custo	muito baixo	baixo	alto	muito alto



Aplicações no Estudo de Compostos Bioativos

Interação Ligante-Receptor: Aspectos Teóricos



$[E_{(aq)}]$

$[I_{(aq)}]$

$[E \cdot I_{(aq)}]$

$$K_i = K_d = \frac{[E_{(aq)}] [I_{(aq)}]}{[E \cdot I_{(aq)}]}$$

$$\Delta G_{bind} = -RT \ln K_a = RT \ln K_d$$

$$\Delta G_{bind} = \Delta H_{bind} - T \Delta S_{bind}$$

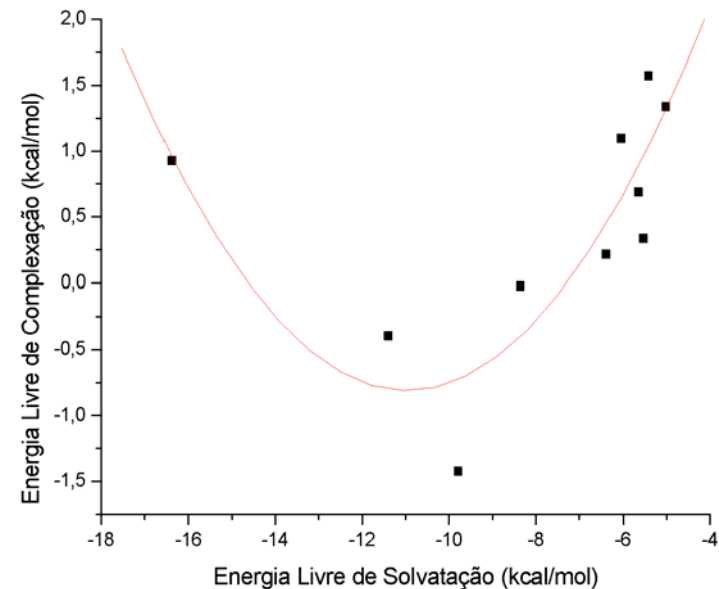
Trocas de energias envolvidas no processo de "binding", como as energias de interação ligante-sítio e ligante-solvente.

Formas de distribuição das energias, classificadas como entropia translacional, rotacional, conformacional e vibracional.

Atividade Biológica: Construção de Modelos Teóricos

Objetivo: proposição de um modelo teórico capaz de determinar a atividade biológica.

$$\Delta G = RT \ln K_i$$



$$RT \ln K_i = c_1 (\Delta G_{solv} + c_2)^2 + c_3 \Delta H_{bind} + c_4 N_{LR} + c_5$$

$$\Delta H_{bind} = \Delta H_{comp} - [\Delta H_{sa} + \Delta H_{lig}]$$

N_{LR} : termo entrópico associado à perda de rotação em ligações após associação do ligante ao sítio ativo

Referências

- Harel, M., Kryger, G., Rosenberry, T. L., Mallender, W. D., Lewis, T., Fletcher, R. J., Guss, J. M., Silman, I., Sussman, J. L. (2000) *Protein Science* **9**, 1063.
- Liverton, N. J.; Butcher, J. W.; Claiborne, C. F. et al. (1999) *J Med Chem.* **42**, 2180.
- Mutero, A., Pralavorio, M., Bride, J., Fournier, D. (1994) *Prod. Natl. Acad. Sci. USA* **91** 5922.
- Ott, K. et al., (1996) *J. Mol Biol.*, **263**, 329.
- Sant'Anna, C. M. R. e Santos, A. C. S. (2002) *J. Mol. Struc. Theochem*, 585, 59.
- Sant'Anna, C. M. R., Souza, V. P., Andrade, D. S. (2002) *Int. J. Quantum Chem.* **87**, 311.
- Sant'Anna, C. M. R., Viana A. S. Nascimento Junior, N. M. (2006) *Bioorg. Chem.* 34, 77.
- Santos, V. M. R., Costa, J. B. N., Moya, G. E. M., Sant'Anna, C. M. R., (2004) *Phosph. Sulf. Sil. Rel. Elem.* **179**, 173.
- Santos, V. M. R., Costa, J. B. N., Moya, G. E. M., Cortes, W. S. e Sant'Anna, C. M. R., (2007) *Bioorg. Chem.* 35.

Referências

- Silva, G. M. S., Sant'Anna, C. M. R., Barreiro, E. J. (2004) *Bioorg Med Chem Lett.* **2004**, 12, 3159.
- Stewart, J. J. P. *J. Comp. Chem.* **10** (1989) 209.
- Walsh, S. B., Dolden, T. A., Moores, G. D., Kristensen, M., Lewis, T., Devonshire, A. L., Williamson, M. S. *Biochem. J.* **359** (2001) 175.
- Stelmach, J. E.; Liu, L.; Patel, S. B.; et al. *Bioorg Med Chem Lett.* **2003**, 13, 277.

Agradecimentos

CNPq, Faperj

Dr. João Batista N. Costa
(PQ/UFRuralRJ)

Dr. Marco Edilson F. Lima
(PQ/UFRuralRJ)

Dra. Ana Cristina S. Santos
(PQ/UFRuralRJ)

Dr. Gonzalo E. Moya
(PQ/PUFRuralRJ)

Dr. Wellington S. Cortes
(PQ/UFRuralRJ)

Dr. Eliezer Barreiro
(PQ/LASSBio/UFRJ)

Anivaldo X. da Silva
(PG/UFRuralRJ)

Catarina del Cistia (PG/UFRuralRJ)

Janaina Marques (PG/UFRuralRJ)

Miguel R. F. Neto (PG/UFRuralRJ)

Nailton M. N. Júnior (IC/UFRuralRJ)

Fernanda Guedes (PG/IQ e
LASSBio/UFRJ)

