# Energy minimization methods and drug design

Yeona Kang
Applied Mathematics and Statistics

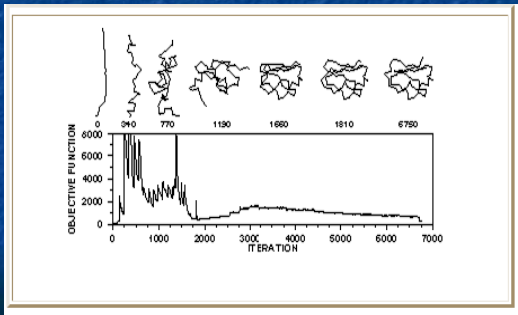
1

INTRODUCTION

- 1. Energy minimization for protein folding
- 2. several minimization methods
  - Monte Carlo Method
  - Steepest Descent Method
  - Conjugate Gradient Method
- 3. Example for this method –drug design
- 4. Conclusion

2

## Energy Minimization




3

## Energy minimization

- Removes atomic overlaps and unnatural strains in the structure
- Stabilizes or reinforces strong hydrogen bonds, breaks weak ones

4

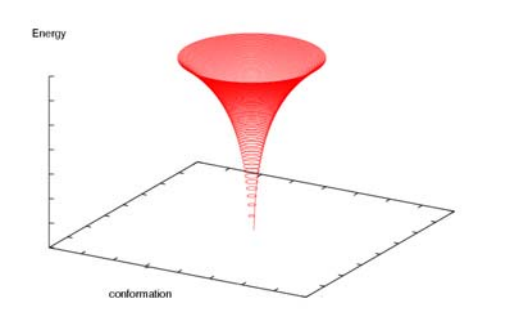
## Standard Energy Function

$$E = K_r(r_i - r_j)^2 + K_\theta(\theta_i - \theta_j)^2 + K_\phi(1-\cos(n\phi_j))^2 + q_iq_j/4\pi\varepsilon r_{ij} + A_{ij}/r^6 - B_{ij}/r^{12} + C_{ij}/r^{10} - D_{ij}/r^{12}$$

Bond length
Bond bending
Bond torsion
Coulomb
van der Waals
H-bond

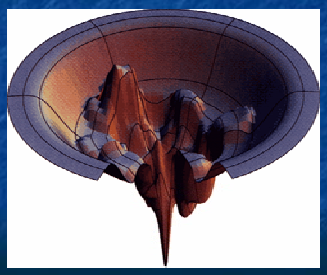
5

## An Energy Surface




6

## A More Realistic Protein Energy Surface



Folding funnel

7

## Minimization Methods

- Energy surfaces for proteins are complex hyperdimensional spaces
- Biggest problem is overcoming local minimum problem
- Simple methods (slow) to complex methods (fast)
  - Monte Carlo Method
  - Steepest Descent
  - Conjugate Gradient

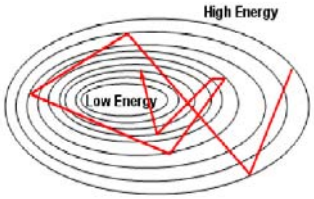
8

## Monte Carlo Algorithm

- Generate a conformation or alignment (a state)
- Calculate that state's energy or "score"
- If that state's energy is less than the previous state
  accept that state and go back to step 1
- If that state's energy is greater than the previous state
  accept it if a randomly chosen number is $< e^{-E/kT}$ where E is the state energy otherwise reject it
- Go back to step 1 and repeat until done

9

## Monte Carlo Minimization




Performs a progressive or directed random search

## Steepest Descent & Conjugate Gradient

- Frequently used for energy minimization of large (and small) molecules
- Ideal for calculating minima for complex (I.e. non-linear) surfaces or functions
- Both use derivatives to calculate the slope and direction of the optimization path
- Both require that the energy function be differentiable (smooth)

11

## Steepest Descent - Conjugate Gradient Method

$\frac{d^2E}{dx^2} = \gamma$ : universal constant –step size

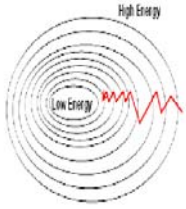
$x_0 - x_1 = -\gamma \frac{dE}{dx}$ : estimation equation of the distance from $x_1$ to the minimum $x_0$

In steepest descents method,
both the gradient and the direction of successive steps are orthogonal.

In conjugate gradient,
the gradient at each point are orthogonal but the directions are conjugate

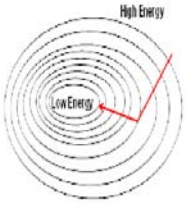
12

## Steepest Descent Minimization




Makes small locally steep moves down gradient

## Conjugate Gradient Minimization

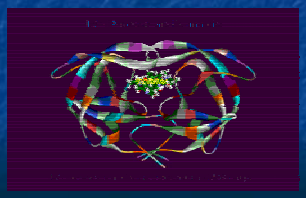



Includes information about the prior history of path

13

## Drug Design Using the Energy Minimization

Numerous drugs are enzyme inhibitors which may be divided into two types: **competitive inhibitors** and **non-competitive inhibitors.**
The competitive inhibitor directly occupies the active site before the substrate can bind to it. The non-competitive inhibitor does not occupy the active site, but its binding to other region somehow inhibits the catalytic activity.
HIV protease inhibitor is a competitive inhibitor



14

## Drug Design Using the Energy Minimization

- The simple energy minimization of the proposed inhibitors in the enzyme active site leads to an intermolecular energy which correlates highly with enzyme inhibition.
- Given two inhibited enzyme:

  The native, acetylpepstatin-inhibited, and L-689,502-inhibited HIV-1 protease X-ray coordinates

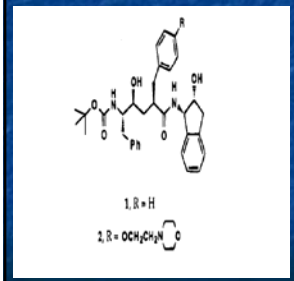
15

## Assumption

- A model of 1 : constructed in the Merck molecular modeling program AMF based on the X-ray structures of inhibitors of endothiapepsin and Rhizopus pepsin.
- Models of all other inhibitor : employing the model of 1 as a template.
- All inhibitors are neutral (no ionic charges).
- Flexibility : necessary to obtain a satisfactory fit in the enzyme active site, which also corresponded to a low-energy conformer.

16

## A model of 1




- The HIV-1 protease inhibitor1 be optimized from initial inhibitor lead and as a small subnanomolar inhibitor.
- Use of this model led to the successful qualitative design of and improved inhibitor.

17

## Energy minimization

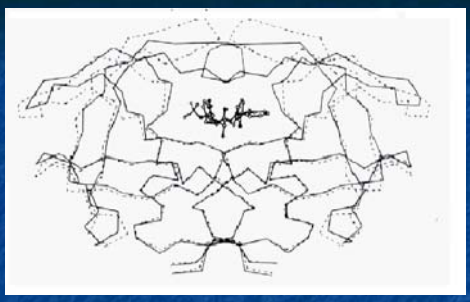
- Inhibitor models was minimized in the three enzyme active sites using the MM2X force field implemented in the program OPTIMOL.
- The final inhibitor model is chosen on the basis of the lowest total energy; the balance between favorable intermolecular and intramolecular energies.

18

- $E_{inter}$ corresponds to the sum of the van der waals and electrostatic interactions between the inhibitor and the enzyme when the inhibitor is minimized in the rigid enzyme active site.

$$E_{\text{inter}} = E_{\text{vdw}} + E_{\text{elec}}$$

19



A comparison of the X-ray coordinates of the native enzyme (dashed) and the complex (solid) with L-689,502,2 (ball and stick). The two enzymes are represented as C_alpha traces to illustrate the difference in position of the flops( at the top of the figure ) which in the native enzyme X-ray structure are too distant to interact with the inhibitor.

20

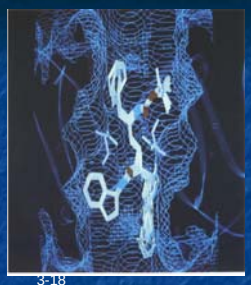
**The correlation seems to be relatively independent of the position of the flaps.**

- Why? Using this paper:

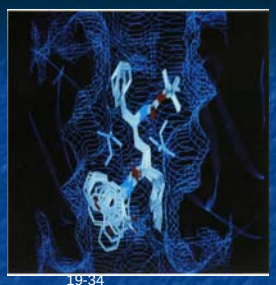
(1) there are specific hydrogen bonds between the inhibitor and the floor of the active site which force the inhibitor to adopt an appropriate bioactive conformation in the absence of the flaps.

(2) The binding elements contributed by the flaps are relatively constant for this series of inhibitors.

21



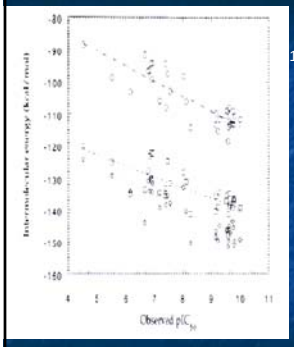
3-18



19-34

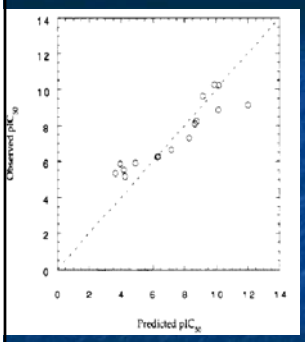
Overlaid models of the training set inhibitors 1 and 3-18, which contain modifications of the $P_1$ substituent, illustrated in the native enzyme active site.

22



1. Plot of calculated enzyme- inhibitor interaction energy vs experimental enzyme inhibition for the training set of inhibitors 1,3-34 .
circles = native enzyme active site
squares = acetylpepstatin-inhibited active site
diamonds = L-689,502-inhibited active site

23




- We can make predictions of activity for many proposed inhibitors prior synthesis.
- The relative accuracy of the these predictions is illustrated graphically in which the line is one of the unit slope,

$$\text{predicted } pIC_{50} = \text{observed } pIC_{50}.$$

24

## *Conclusion*

- Energy minimization
- Several methods of energy minimization
- Application of energy minimization ; Drug design
  - -A prior prediction of activity for HIV-1 protease inhibitors employing energy minimization in the active site.

25

## The analysis of the significant outlier -36

- The factor 3 may play a key role

Why?

" 36 experienced a much larger decrease(13.7kcal/mol) than other inhibitors(2.5kcal/mol) in its intramolecular energy when minimized outside of the active site, an indicate that the bound conformation may be significantly higher in energy than the global minimum."

26

## The analysis of the significant outlier -36

Several possible explanations for overprediction

1. The use of the native enzyme model in eq 1, rather than one of the inhibited enzyme models in eq 2 and 3;
2. The presence of an additional hydrogen bond to the active which would be overemphasized in a gas-phase molecular mechanics calculation;
3. The existence of a higher barrier to obtaining the bioactive conformation necessary for binding.

* The last two seem most likely since the activity predicted using the other two models is also exaggerated.

27

native:

$$pIC_{50} = -0.15435(E_{inter}) - 8.069 \quad (1)$$

$R = 0.8524$, $R^2 = 0.7265$, cross-validated $R^2 = 0.6910$

acetylpepstatin inhibited:

$$pIC_{50} = -0.17302(E_{inter}) - 14.901 \quad (2)$$

$R = 0.7623$, $R^2 = 0.5811$, cross-validated $R^2 = 0.5244$

L-689,502 inhibited:

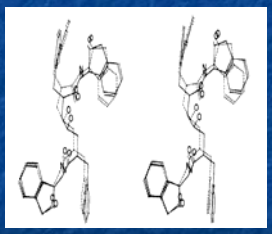
$$pIC_{50} = -0.16946(E_{inter}) - 15.707 \quad (3)$$

Equation 1-3 were derived for the native, acetylpepstatin , and L-689,502 enzyme active sites, respectively.

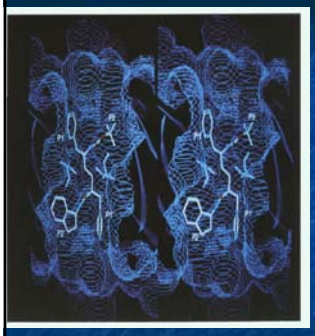
28

## The analysis of the significant outlier -36

- **The significant outlier is 36, which was designed as a symmetrical version of 1.**
- **The activity of 36 be predicted using eq 1.**
- **The predicted activity be substantially exaggerated.**
- **However when the X-ray structure of 36 be solved, there was good agreement between the modeled and X-ray structures.**



29



1. Stereoview of L-685,434,1, as modeled in the native enzyem.
2. The active site cavity is represe-nted by a wire-mesh molecular surface and enzyme is depicted as a ribbon.
3. The catalytic aspartates , Asp_A25 and Asp_B25, are highlighted for Reference.
4. The inhibitor binding elements are labeled following the Schechter and Berger nomencla-ture for protease.

30

We focus on a possible correlation between the calculated interaction energy, the intermolecular component of the total energy, and the observed $IC_{50}$

E corresponds to the sum of the van der waals and Electrostatic interactions between the inhibitor and the enzyme When the inhibitor is minimized in the rigid enzyme active site.

$$E_{\text{inter}} = E_{\text{vdw}} + E_{\text{elec}}$$

Two assumption:

1. E might be proportional to the enthalpy of binding.
2. The entropy of binding might be small or more likely constant.

31