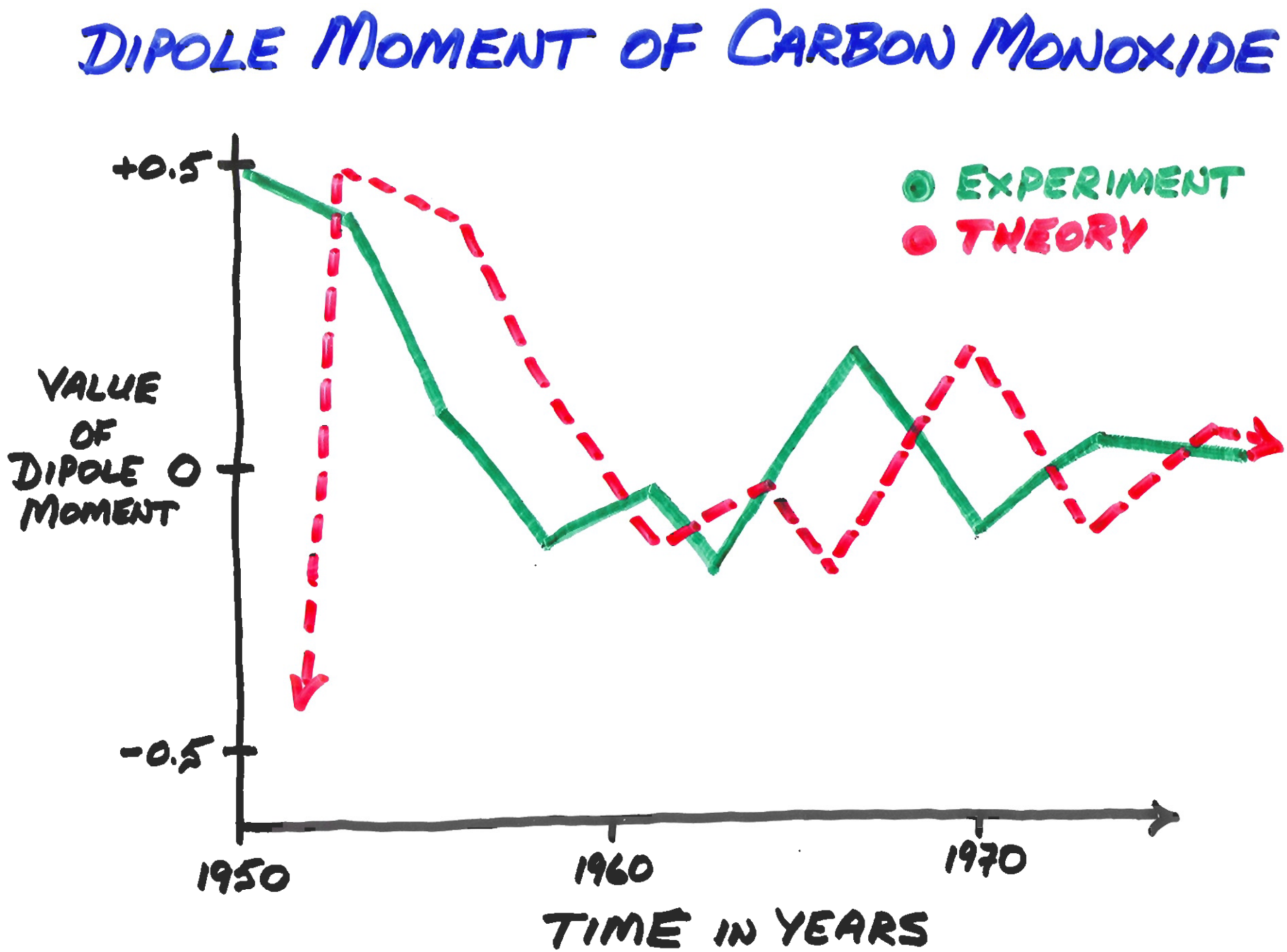
[Blatantly Stolen from Bill Goddard, Computational Chemistry GRC, circa 1986]



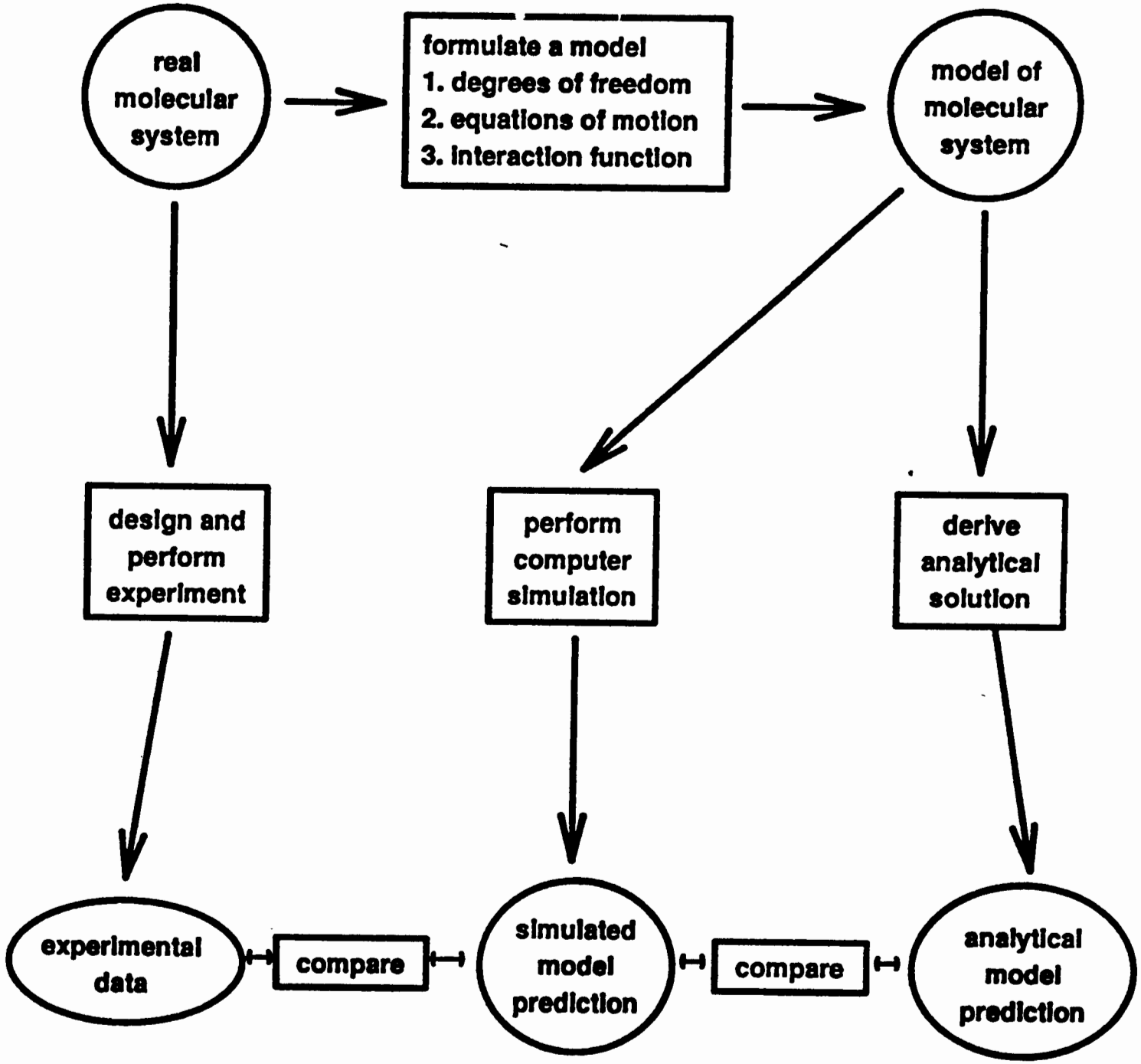


Fig. 1. Molecular models, simulation and experiment.

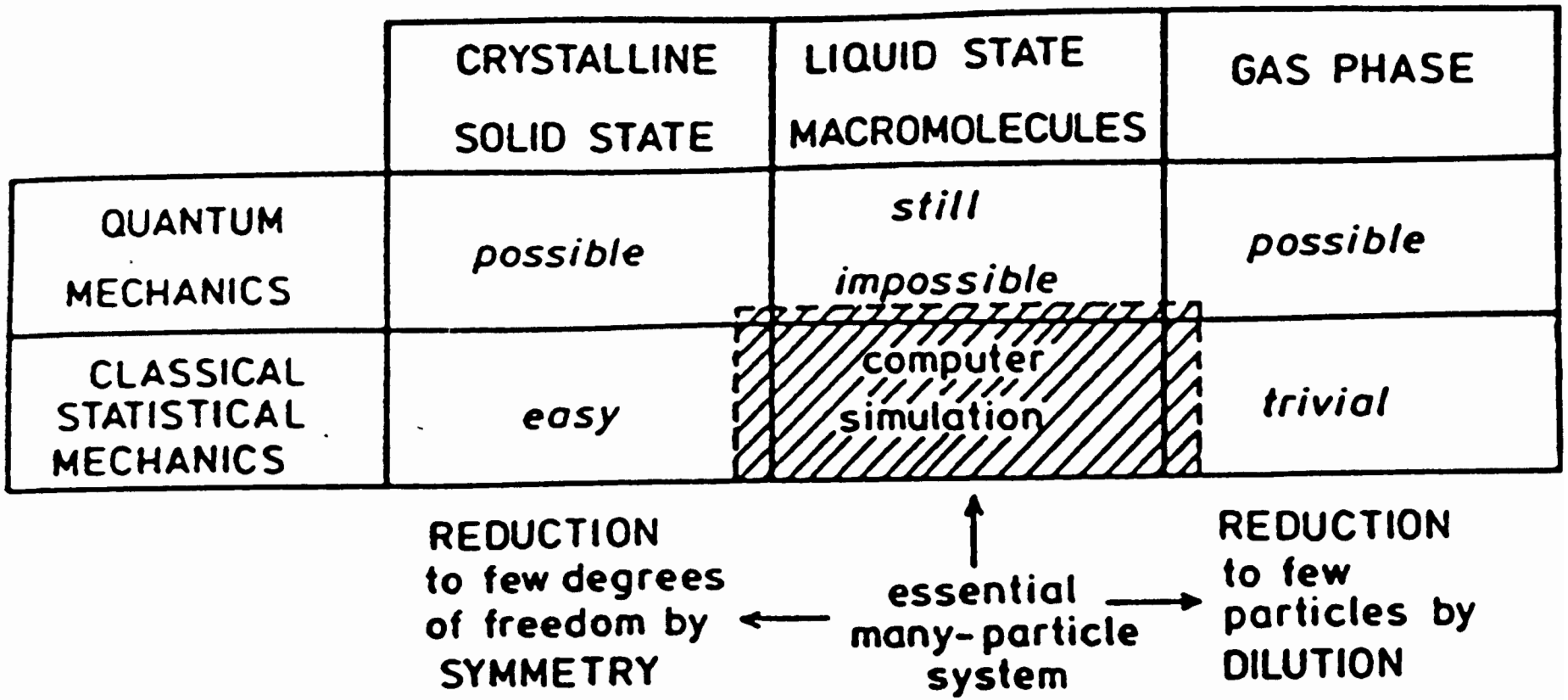


Fig. 1. Classification of molecular systems. Systems in the shaded area are amenable to treatment by computer simulation.

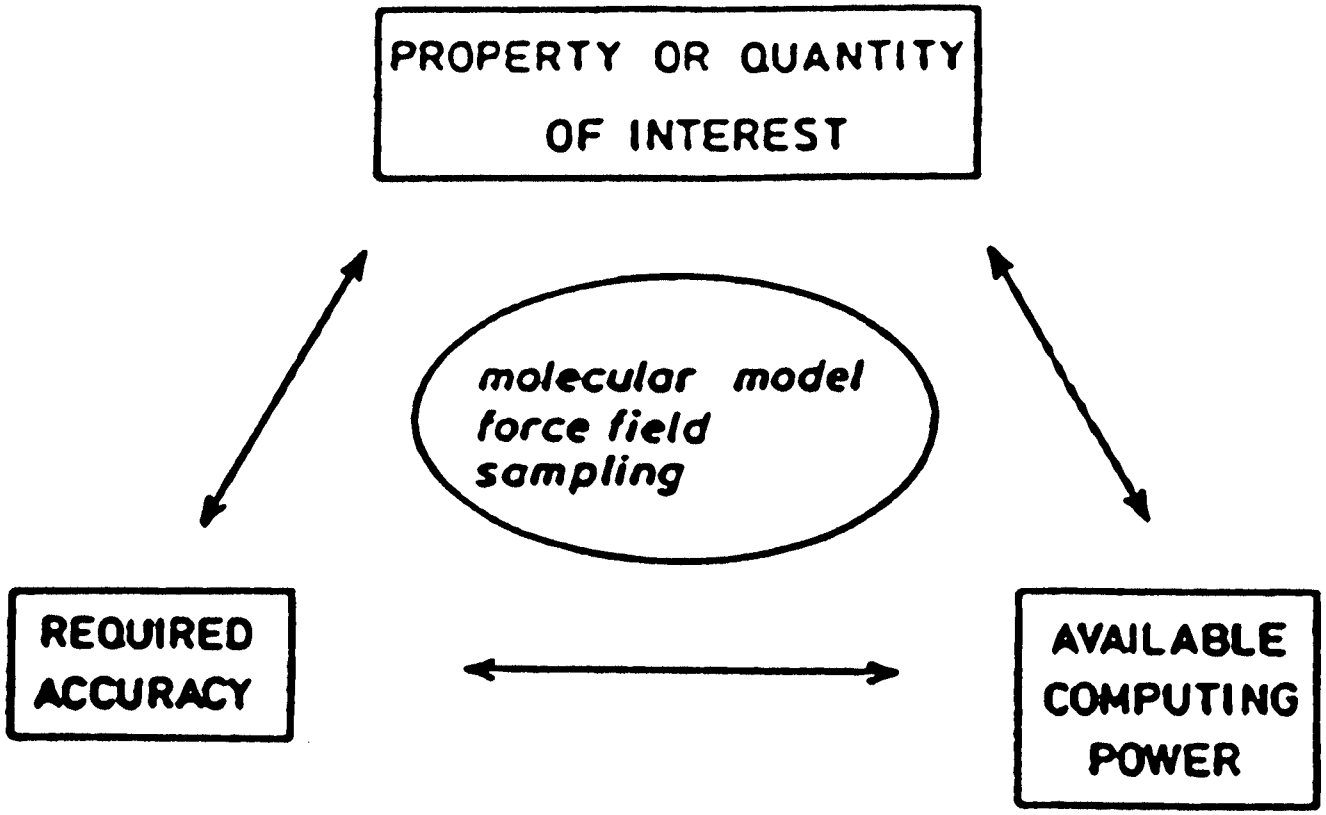


Fig. 3. Choice of molecular model, force field and sample size depends on 1) the property one is interested in (space to be searched), 2) required accuracy of the prediction, 3) the available computing power to generate the ensemble.

Table 3.1. *Typical features of some internal motions of proteins and nucleic acids*

Motion	Spatial extent (nm)	Amplitude (nm)	Log_{10} of characteristic time (s)
Relative vibration of bonded atoms	0.2 to 0.5	0.001 to 0.01	−14 to −13
Longitudinal motions of bases in double helices (nucleic acids)	0.5	0.01	−14 to −13
Lateral motions of bases in double helices (nucleic acids)	0.5	0.1	−13 to −12
Global stretching (nucleic acids)	1 to 30	0.03 to 0.3	−13 to −11
Global twisting (nucleic acids)	1 to 30	0.1 to 1.0	−13 to −11
Elastic vibration of globular region	1 to 2	0.005 to 0.05	−12 to −11
Sugar repuckering (nucleic acids)	0.5	0.2	−12 to −9
Rotation of sidechains at surface (protein)	0.5 to 1	0.5 to 1	−11 to −10
Torsional libration of buried groups	0.5 to 1	0.05	−11 to −9
Relative motion of different globular regions (hinge bending)	1 to 2	0.1 to 0.5	−11 to −7
Global bending (nucleic acids)	10 to 100	5 to 20	−10 to −7
Rotation of medium-sized sidechains in interior (protein)	0.5	0.5	−4 to 0
Allosteric transitions	0.5 to 4	0.1 to 0.5	−5 to 0
Local denaturation	0.5 to 1	0.5 to 1	−5 to +1

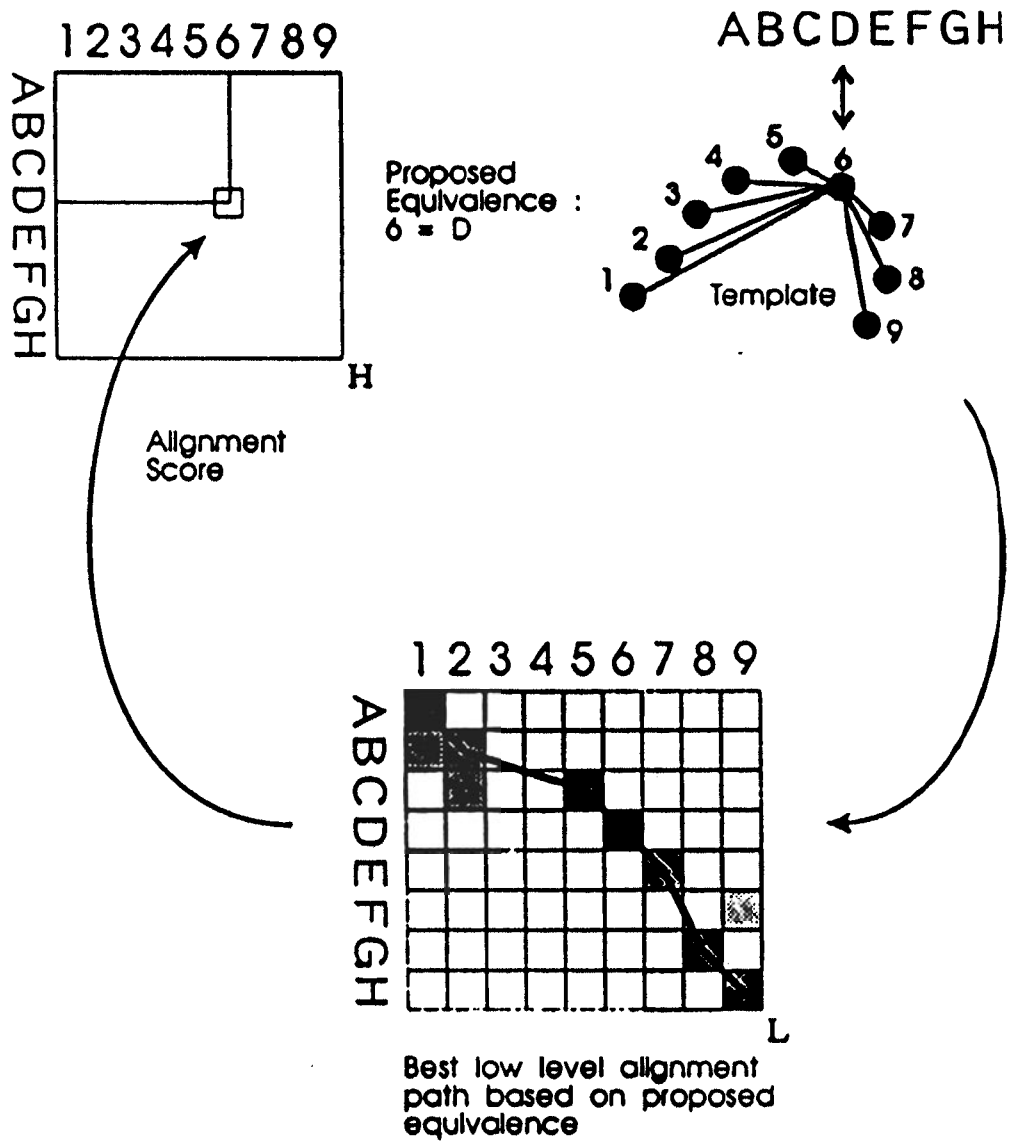
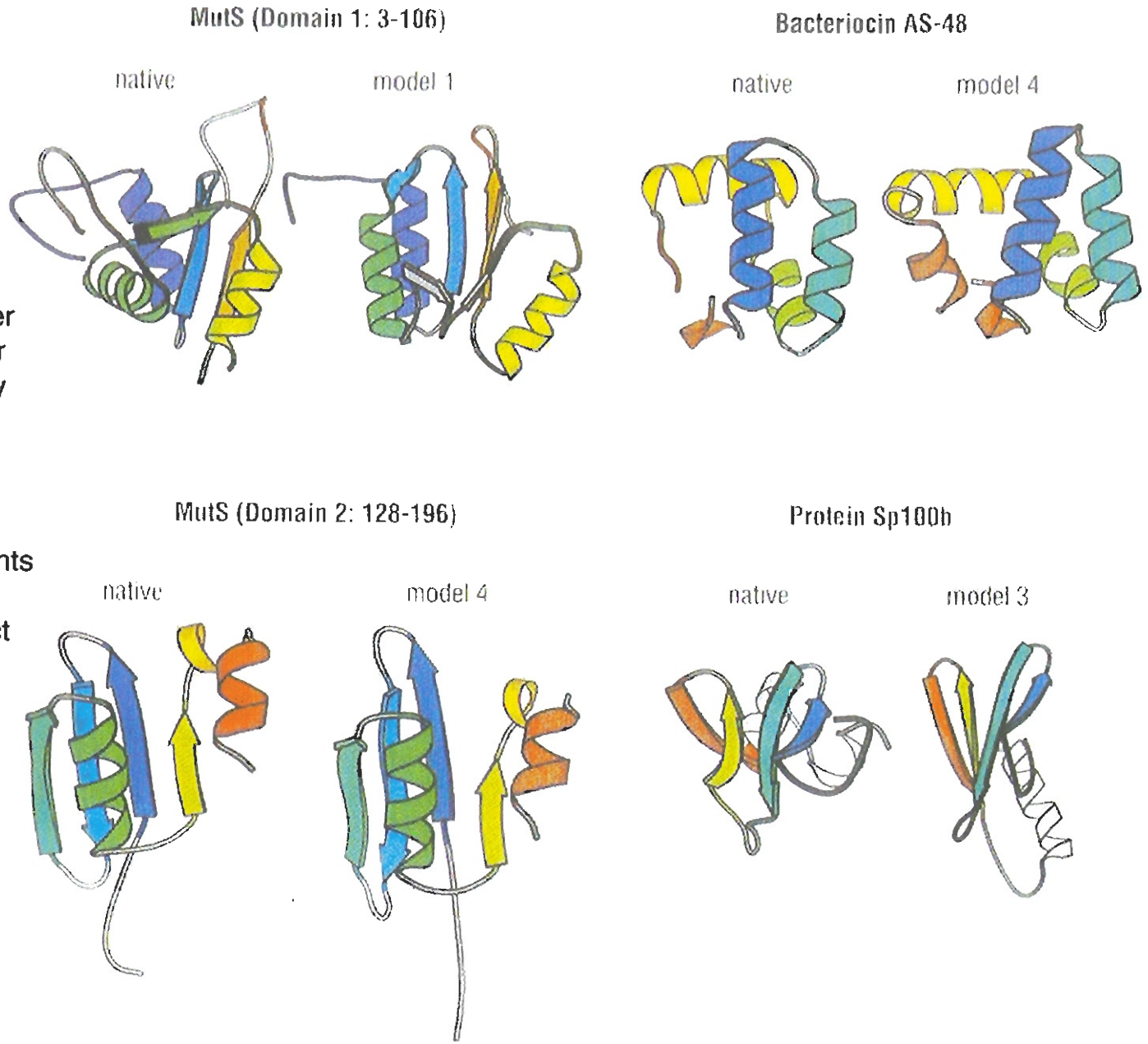


Fig. 4. An illustration of the Double Dynamic Programming algorithm applied to the sequence threading problem. A sequence of amino acids A–H (one letter code) is being threaded onto coordinate positions 1–9 of a structural template. Given the proposed equivalence that residue D lies on position 6, a matrix L of the scores of all other equivalences can be constructed. A best path (or alignment) is then found through this by application of the standard Dynamic Programming algorithm. The overall score for this alignment (which is an indication of how well residue D fits on position 6) is recorded in the matrix H. All potential equivalences are evaluated in this way (filling the matrix H with values) and the best consistent selection of these is found by application of the Dynamic Programming algorithm to the H matrix. This double application of the alignment algorithm at two levels gives rise to the name.

Rosetta Uses a Fragment Library + Monte Carlo Search

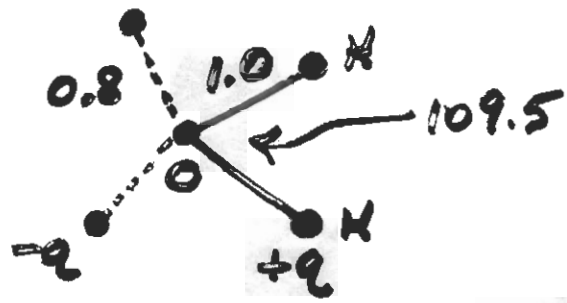
Examples of the best-center cluster found by *Rosetta* for some test proteins. In many cases the overall fold is predicted well enough to be recognizable. However, relative positions of the secondary structure elements are almost always shifted somewhat from their correct values.



EXAMPLE: WATER

STILLINGER (ST2)

$q = 0.2357$



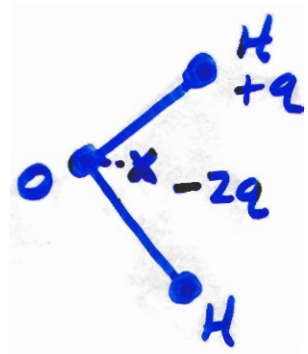
$$E_{LJ} = 0.31 \cdot (\sigma/R)^{12} - 0.31 \cdot (\sigma/R)^{6}, \quad \sigma = 3.10 \text{Å}$$

JORGENSEN (TIP4P)

$O-X = 0.15 \text{Å}$

$q = 0.52$

$R = 0.9572 \quad \theta = 104.52$



$$E_{LJ} = 600/R^{12} - 610/R^{6}$$

(c)

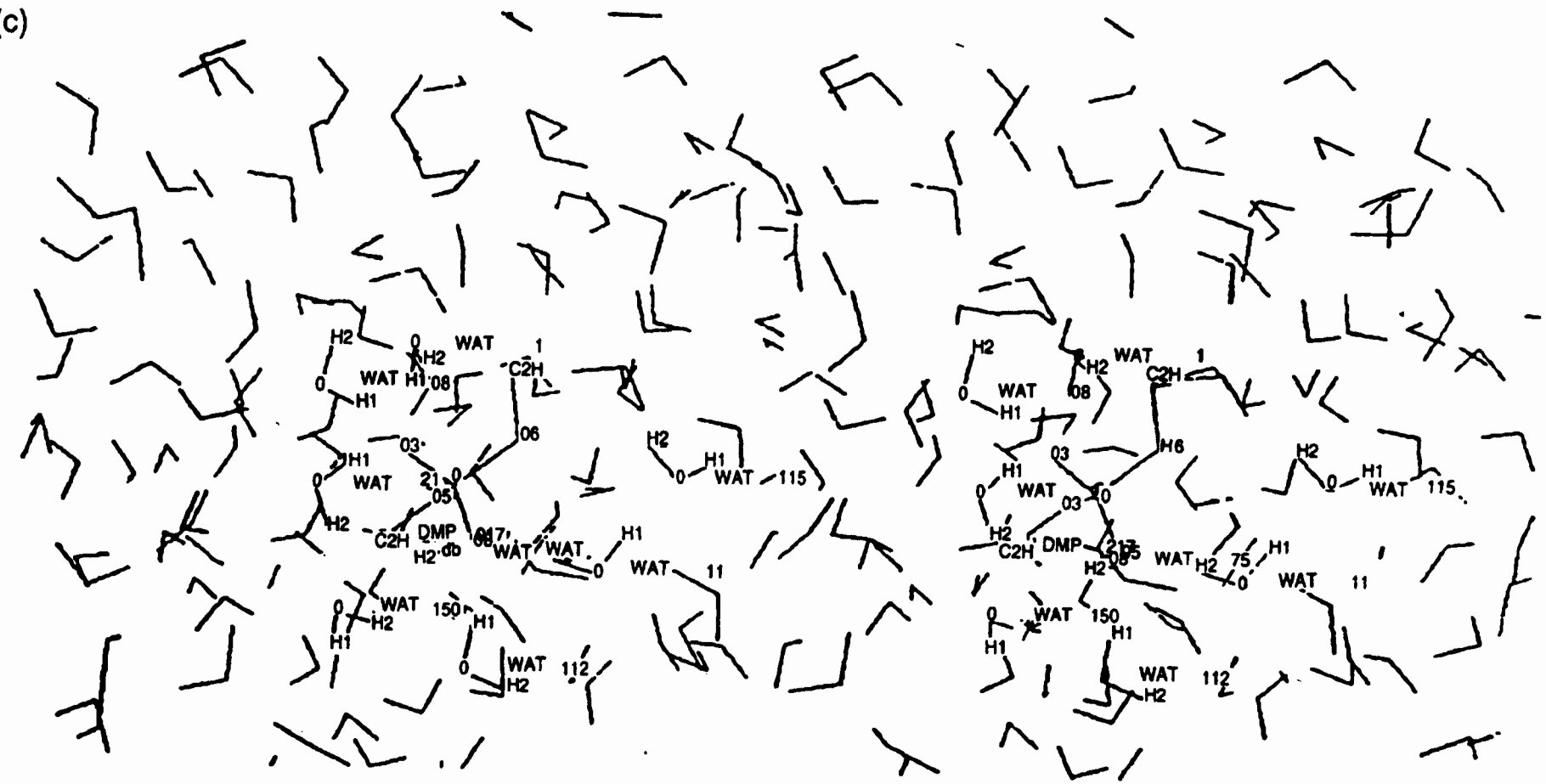


Figure 6.7 (a) Stereoscopic view of the water molecules lying near the anionic oxygens of *g,t* DMP after 5×10^5 steps (top); (b) same as (a) after 7.5×10^5 steps (centre); (c) same as (b) with a different viewpoint and all the water molecules included

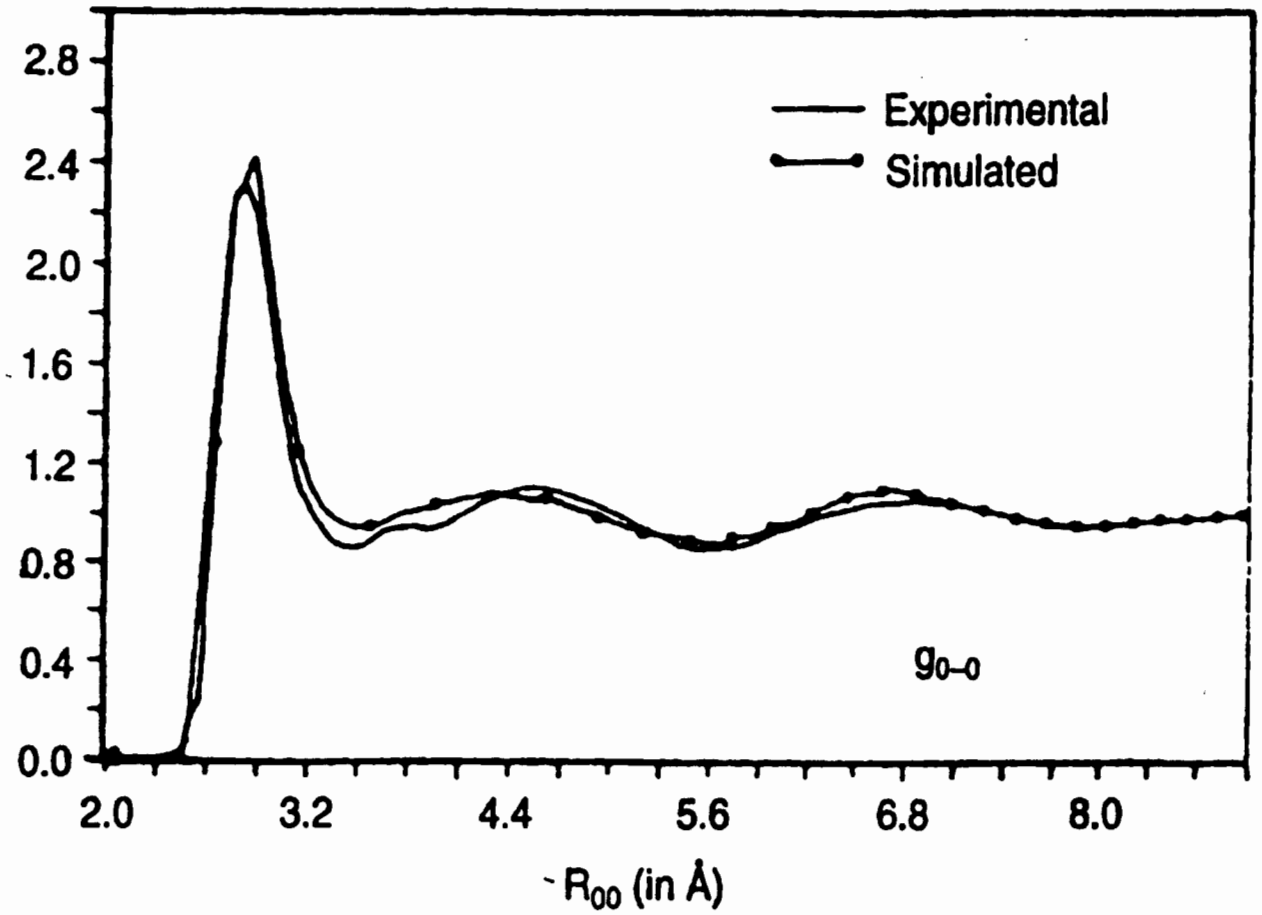


Figure 6.4 Comparison between simulated and experimental O–O radial distribution functions of liquid water (from Lie *et al.* with permission [12]).

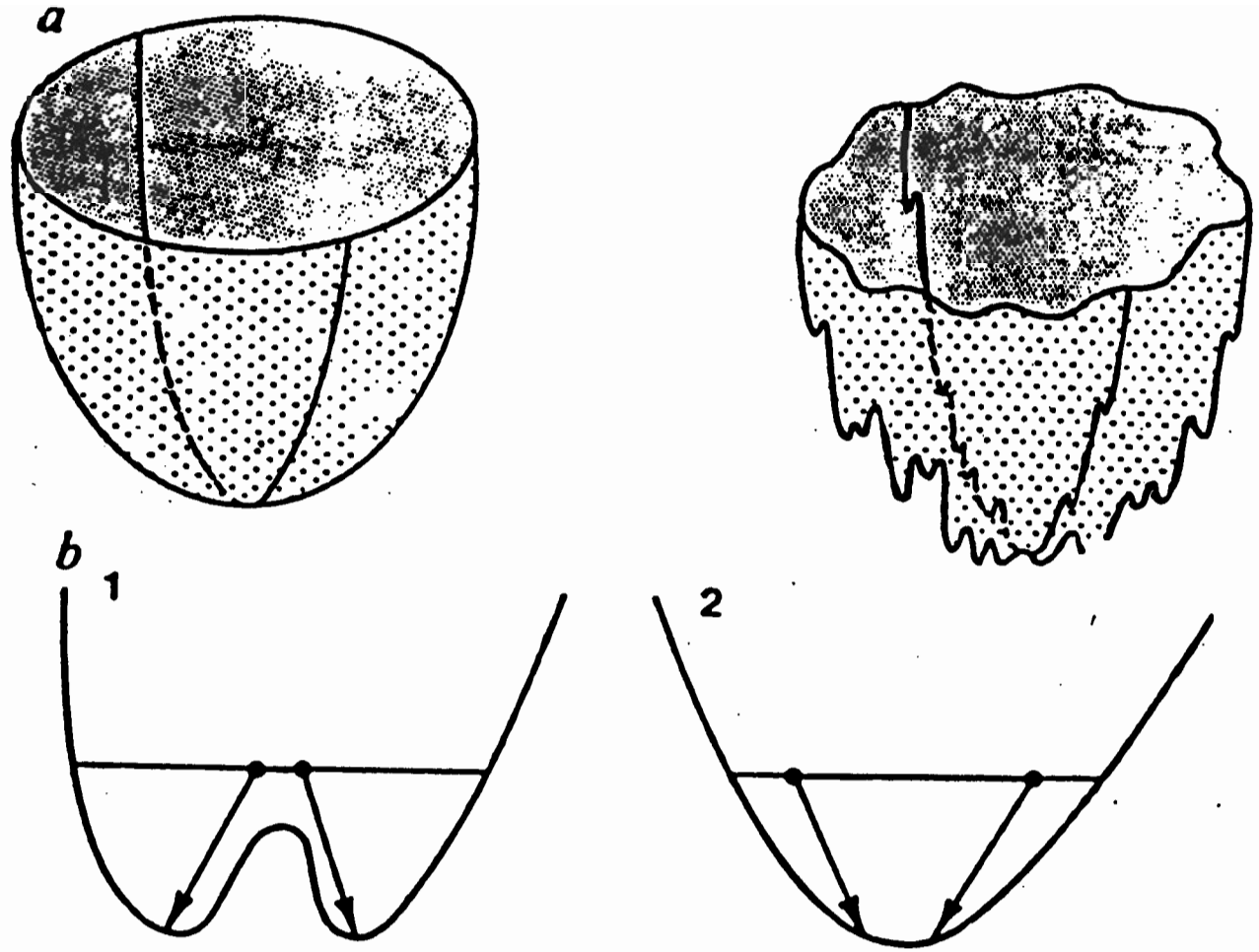
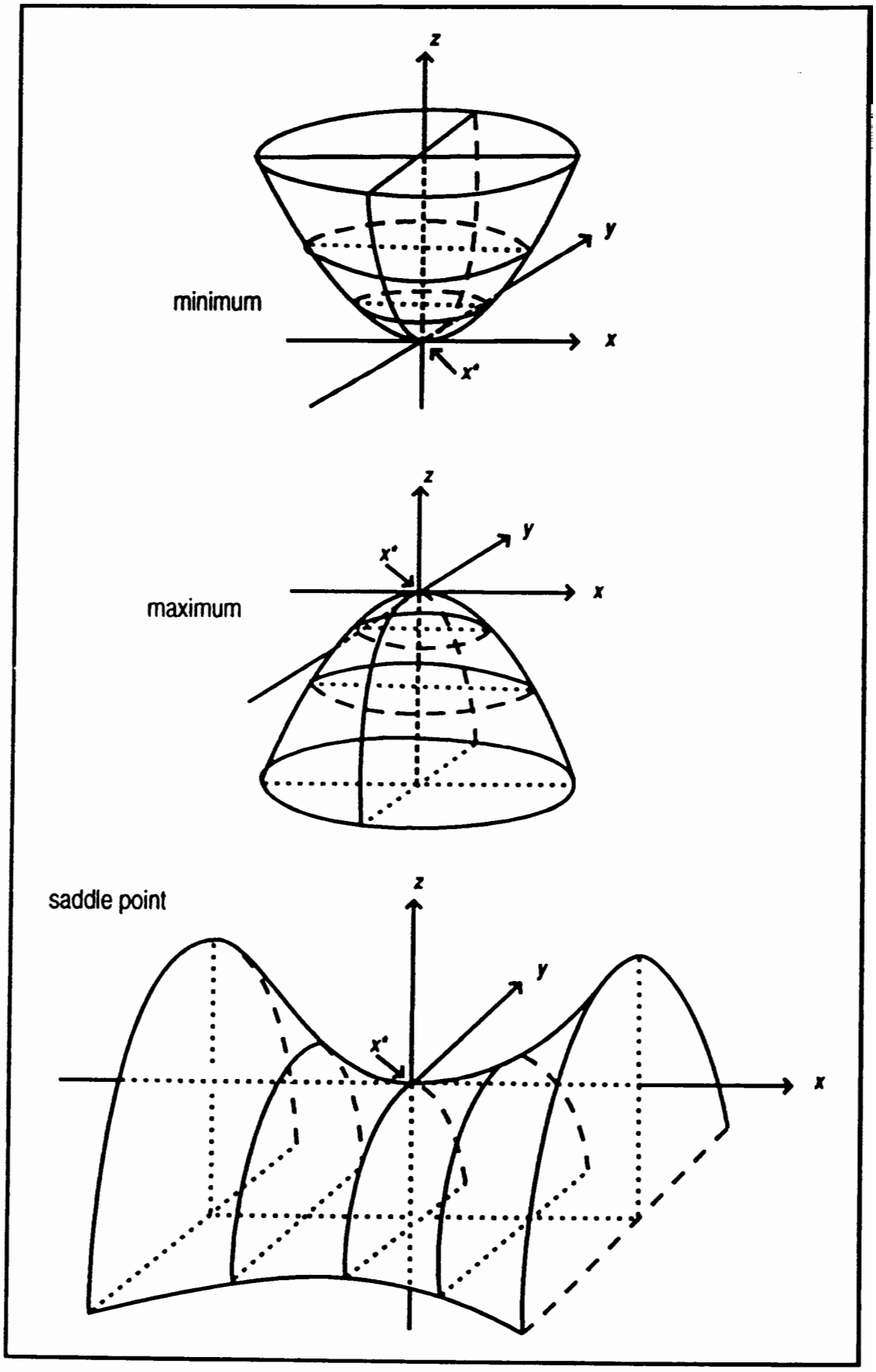


FIG. 3 Model potentials for protein motions. *a*, Two-dimensional representation of harmonic potential (left) and multimimum (substate) potential (right). *b*, representation of the r.m.s. difference criterion for different minima: 1, r.m.s. after the minimization is larger than the initial r.m.s., implying that the two conformations correspond to different minima; 2, r.m.s. after the minimization is smaller than the initial r.m.s., implying that the two conformations correspond to the same minimum (from ref. 13).



Types of stationary points.

```
8  Ethane
1  C       0.000000    0.000000    0.000000    1    2    3    4    5
2  C       0.000000    0.000000    1.530000    1    1    6    7    8
3  H       1.037098    0.000000   -0.366645    5    1
4  H      -0.518502    0.898180   -0.366645    5    1
5  H      -0.518502   -0.898180   -0.366645    5    1
6  H       1.033151    0.090389    1.896645    5    2
7  H      -0.438247   -0.939953    1.896645    5    2
8  H      -0.594811    0.849572    1.896645    5    2


8  Ethane
1  C       1
2  C       1     1   1.53000
3  H       5     1   1.10000     2  109.4700
4  H       5     1   1.10000     2  109.4700     3 -120.0000     0
5  H       5     1   1.10000     2  109.4700     4 -120.0000     0
6  H       5     2   1.10000     1  109.4700     3    5.0000     0
7  H       5     2   1.10000     1  109.4700     3 -115.0000     0
8  H       5     2   1.10000     1  109.4700     3  125.0000     0


8  Ethane
1  C       1
2  C       1     1   1.53000
3  H       5     1   1.10000     2  109.4700
4  H       5     1   1.10000     2  109.4700     3  109.4700     1
5  H       5     1   1.10000     2  109.4700     3  109.4700    -1
6  H       5     2   1.10000     1  109.4700     3    5.0000     0
7  H       5     2   1.10000     1  109.4700     6  109.4700     1
8  H       5     2   1.10000     1  109.4700     6  109.4700    -1
```

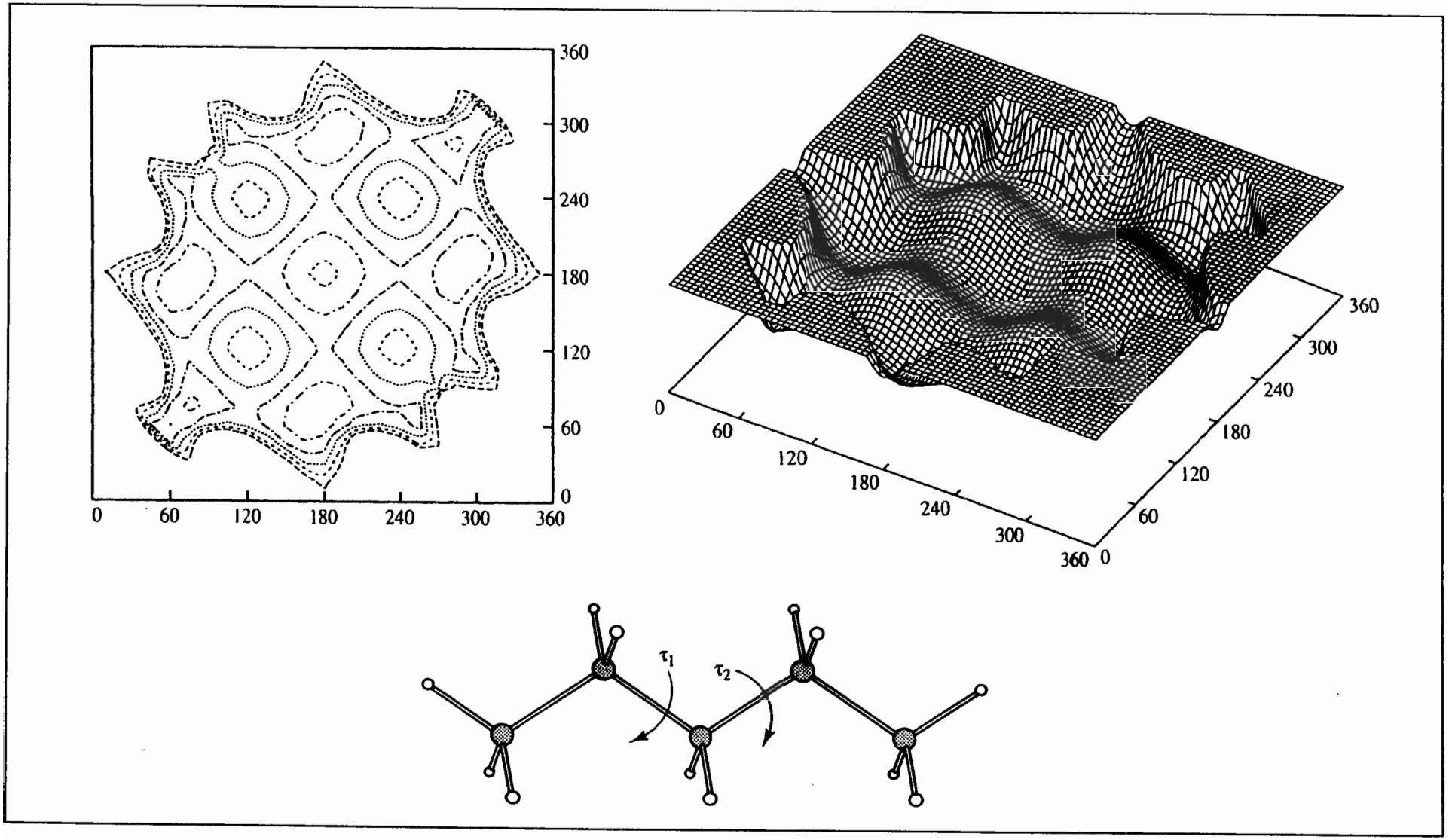
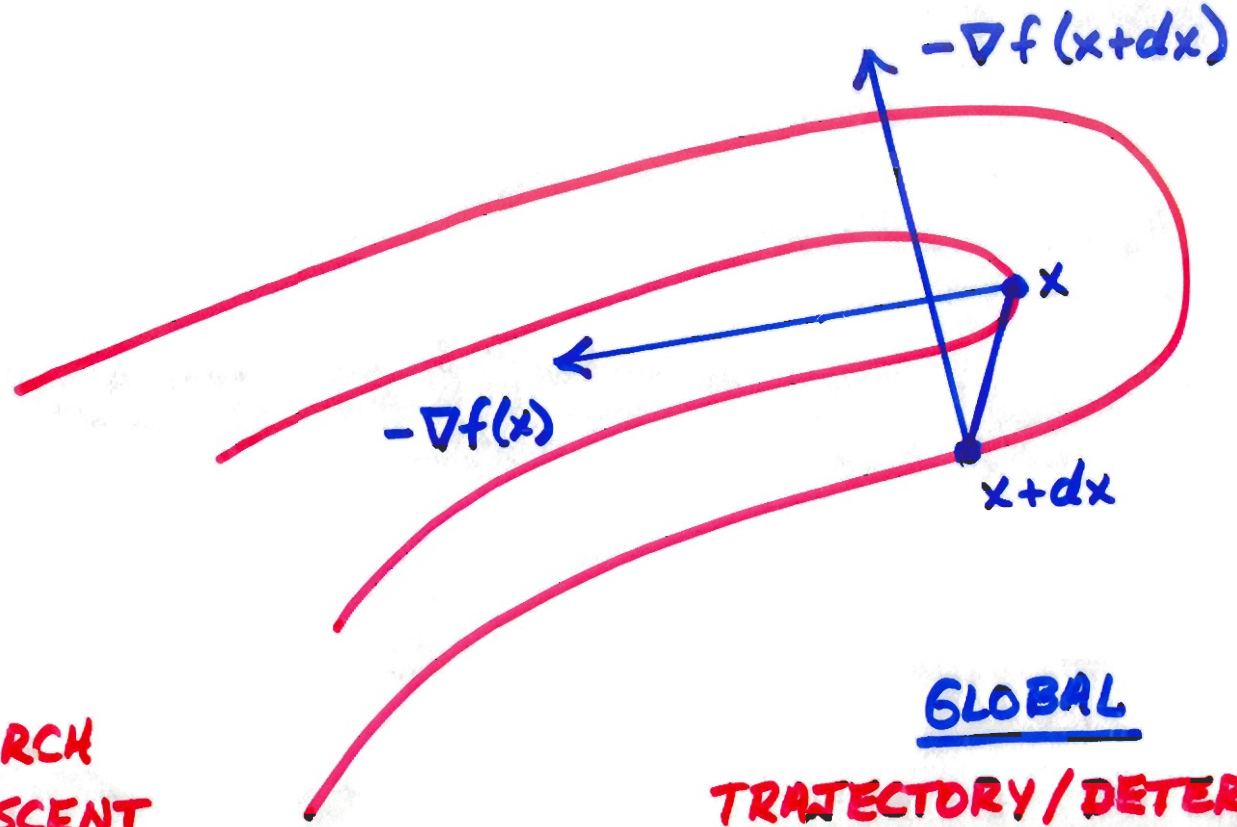


Fig. 5.1: Variation in the energy of pentane with the two torsion angles indicated and represented as a contour diagram and isometric plot. Only the lowest-energy regions are shown.

OPTIMIZATION: LOCAL vs. GLOBAL



LOCAL

PATTERN SEARCH
STEEPEST DESCENT
NEWTON'S METHOD
VARIABLE METRIC

GLOBAL

TRAJECTORY / DETERMINISTIC
BREMMERMAN'S METHOD
CLUSTERING (TORN)
CONTROLLED RANDOM SEARCH
STOCHASTIC METHODS

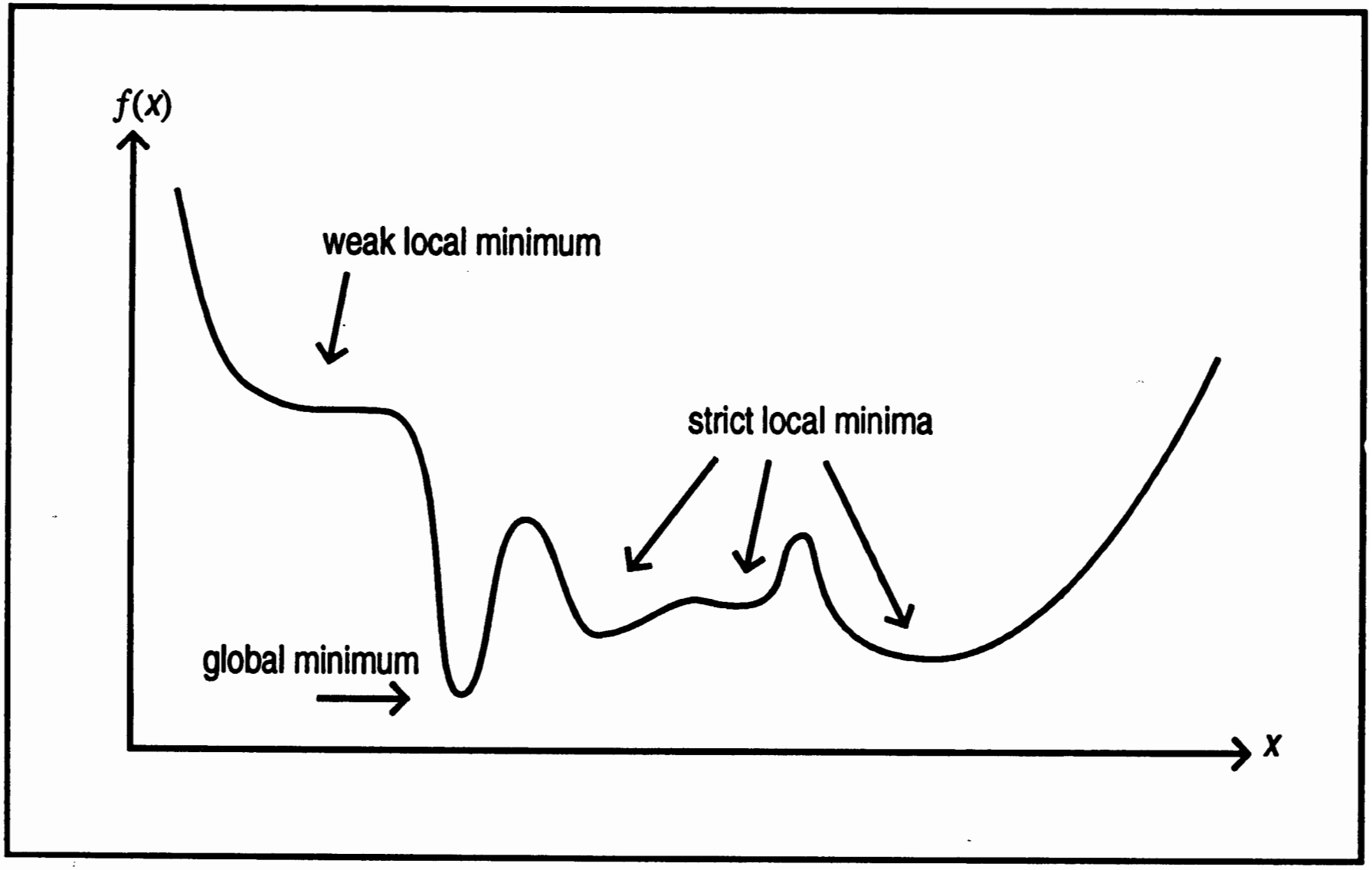
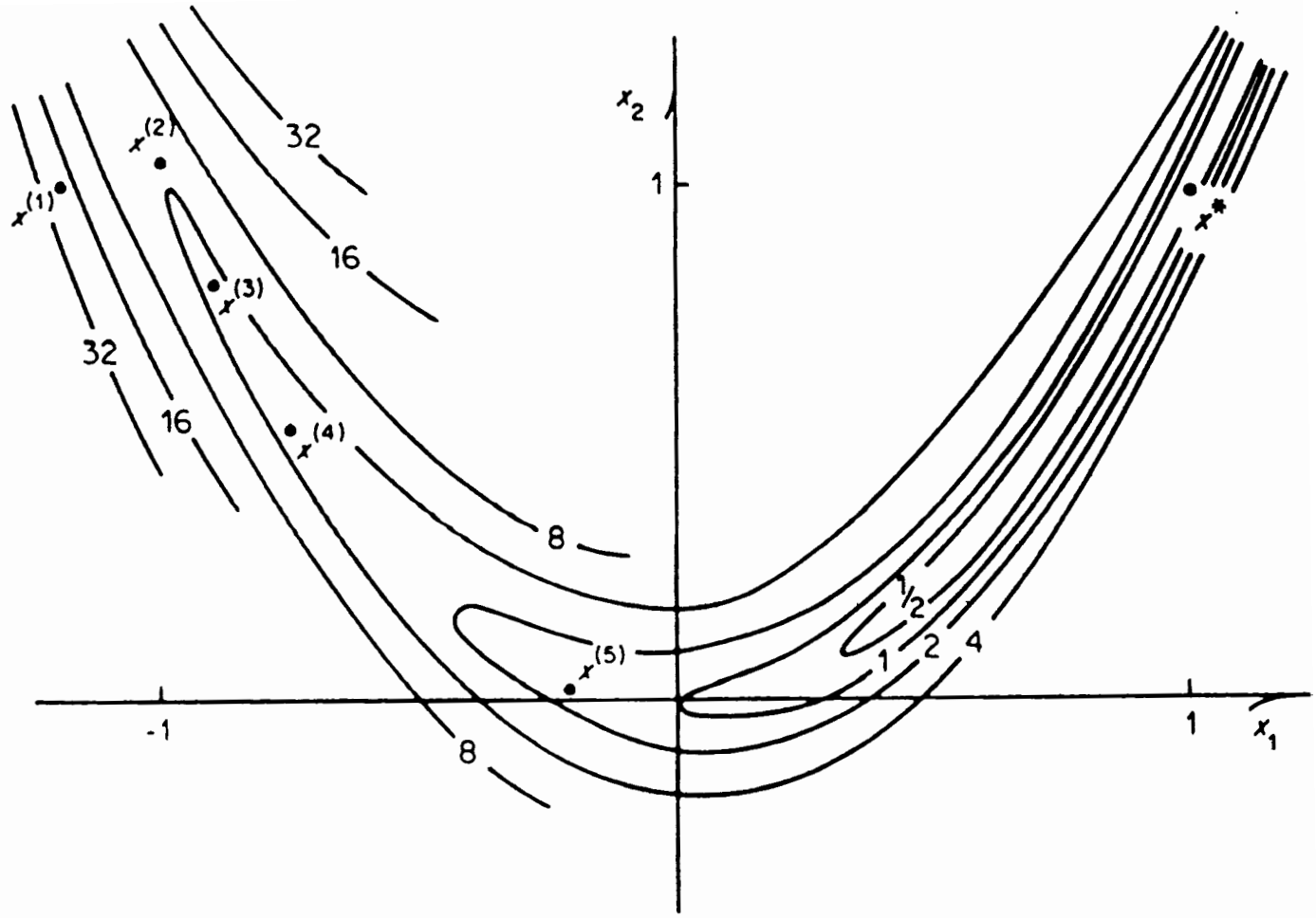


Figure 1 Types of minima.



Contours for Rosenbrock's function

Numerical Example I: Rosenbrock Minimization

Rosenbrock's function is often used as a minimization test problem, because its minimum lies at the base of a "banana-shaped valley" and can be difficult to locate. This function is defined for even integers n as the sum

$$f(\mathbf{x}) = \sum_{j=1,3,5,\ldots,n-1} [(1 - x_j)^2 + 100(x_{j+1} - x_j^2)^2].$$

The contour plot of Rosenbrock's function for $n = 2$ is shown in Figure 14. The minimum point is (1,1), where $f(\mathbf{x}) = 0$. The gradient components of this function are given by

$$\left.\begin{aligned} g_{j+1} &= 200(x_{j+1} + x_j^2) \\ g_j &= -2[x_j g_{j+1} + (1 - x_j)] \end{aligned}\right\}, \qquad j = 1,3,5,\ldots,n-1,$$

and the Hessian is the 2×2 block diagonal matrix with entries

$$\left.\begin{aligned} H_{j+1,j+1} &= 200 \\ H_{j+1,j} &= -400x_j \\ H_{j,j} &= -2(x_j H_{j+1,j} + g_{j+1} - 1) \end{aligned}\right\}, \qquad j = 1,3,5,\ldots,n-1.$$

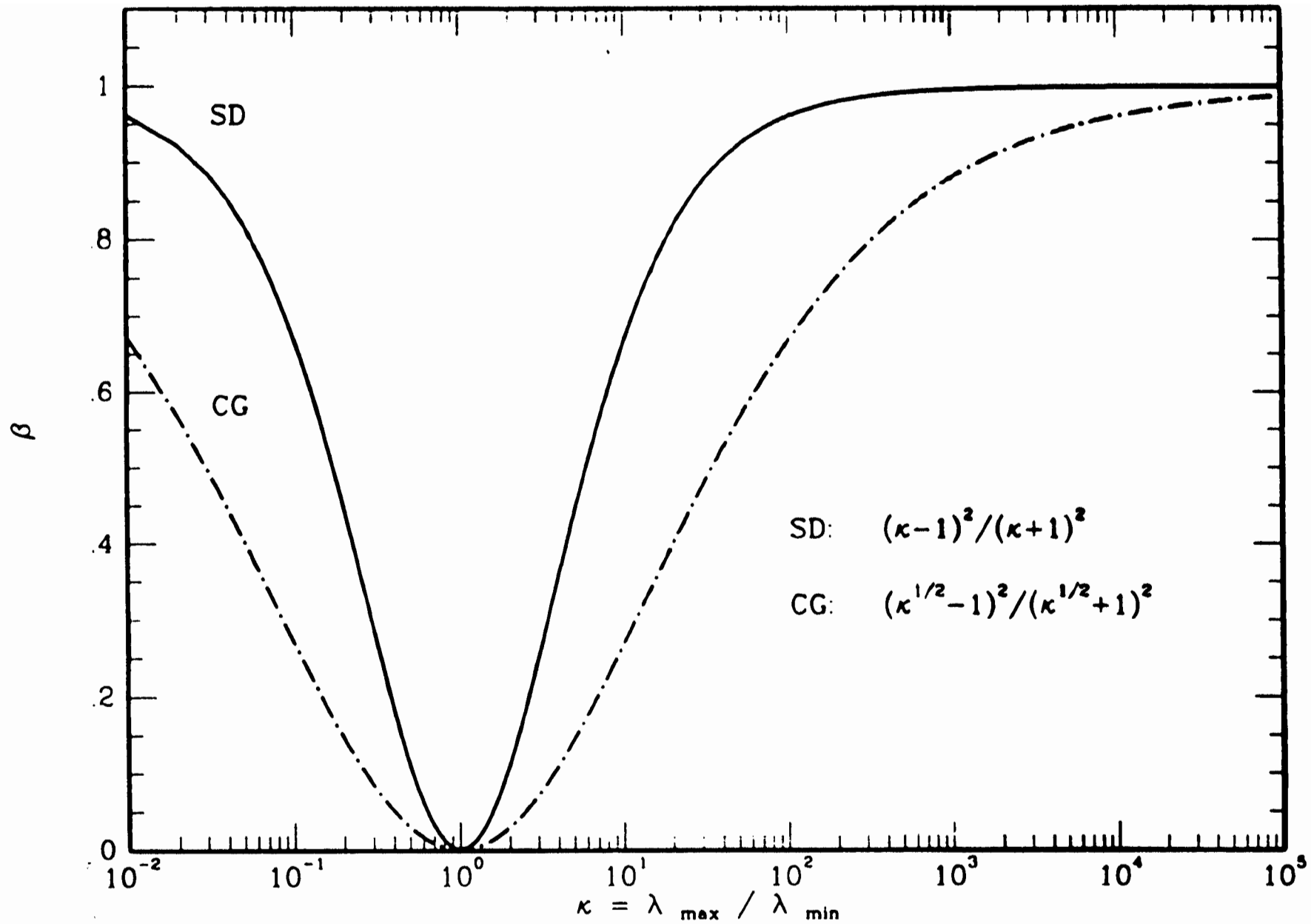
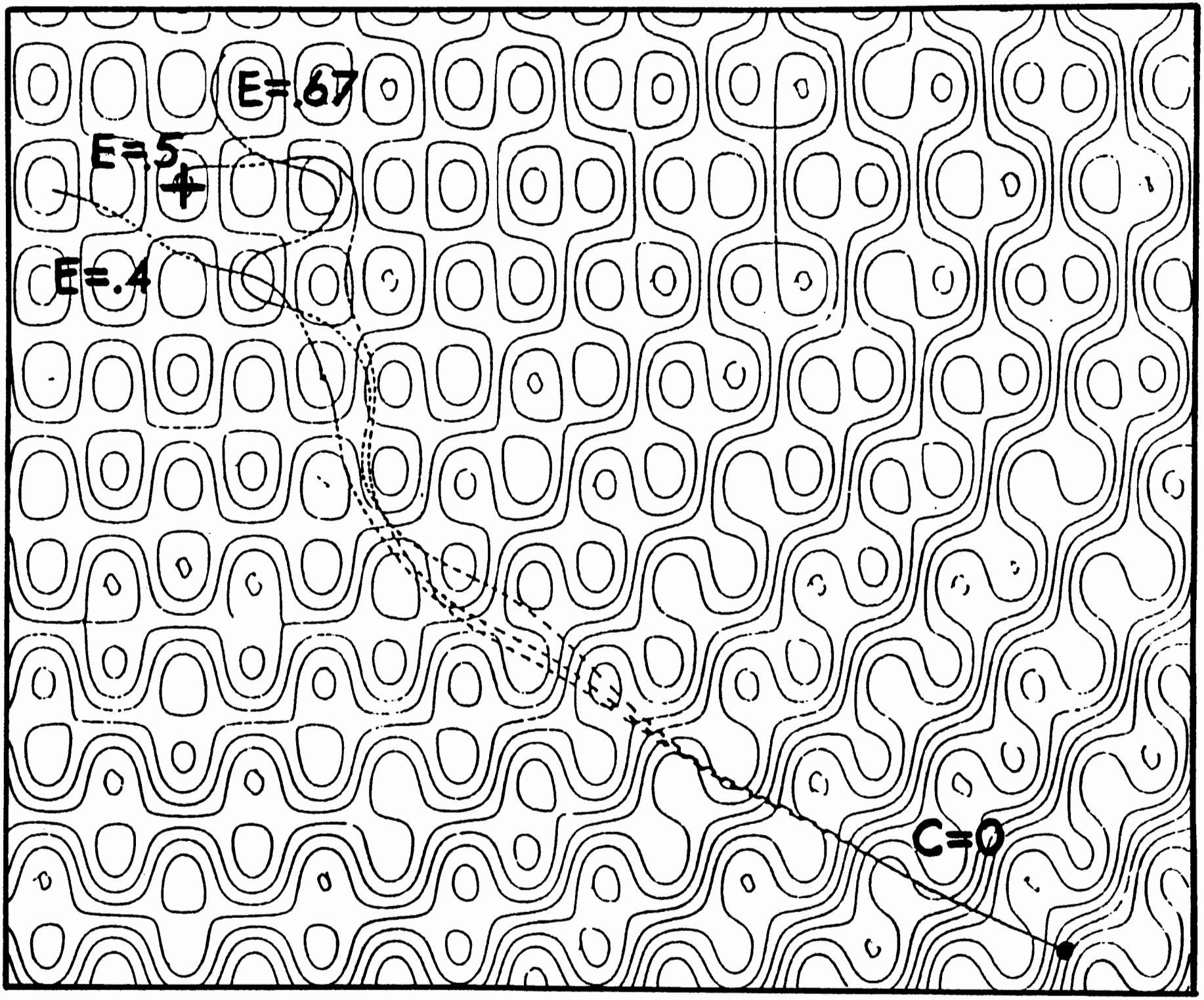
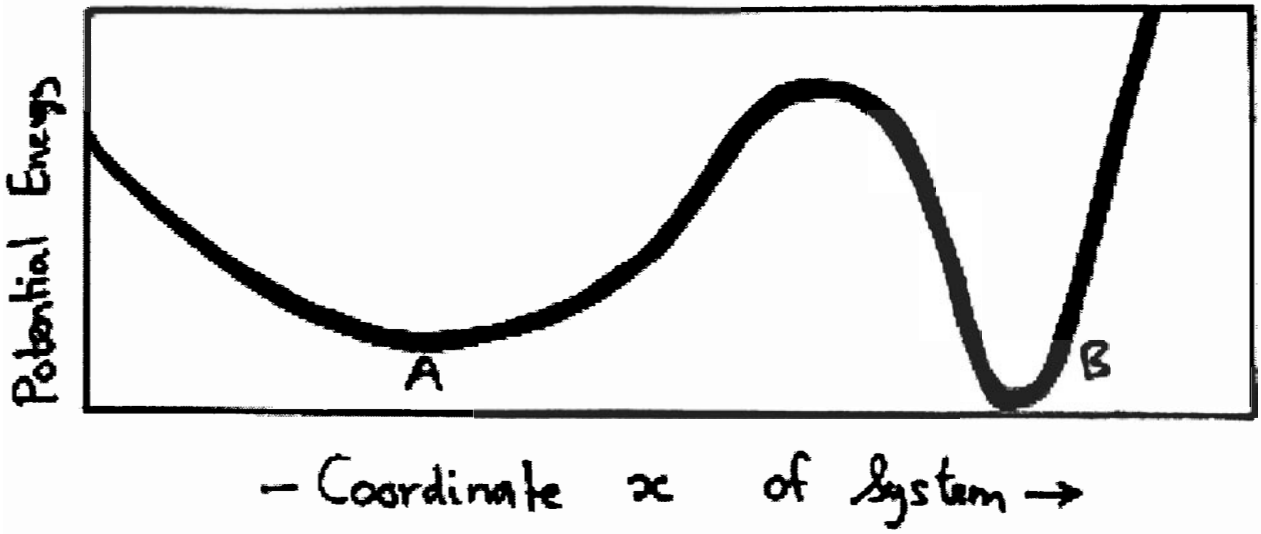


Figure 12 Steepest descent and conjugate gradient quantities that affect the convergence rate for quadratic functions (see text for the distinct context of these functions).


E=.67
E=.5
E=.4
C=0

BOLTZMANN'S DISTRIBUTION

Imagine a potential energy function $U(x)$.



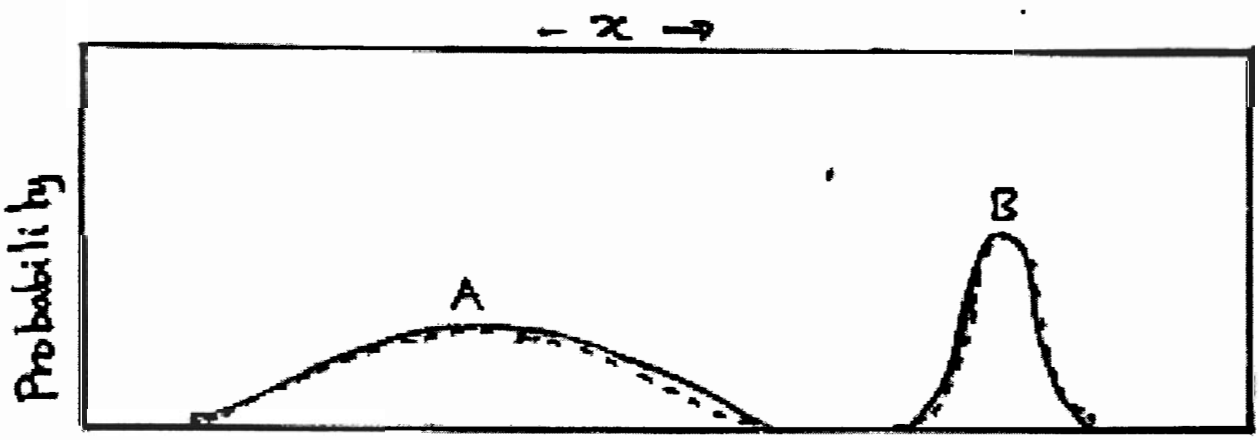
Probability of being at position x is

$$P(x) = \exp(-U(x)/kT) \,/\, (\text{normalization}).$$

normalize so that $\sum_i P(x_i) = 1$ ie

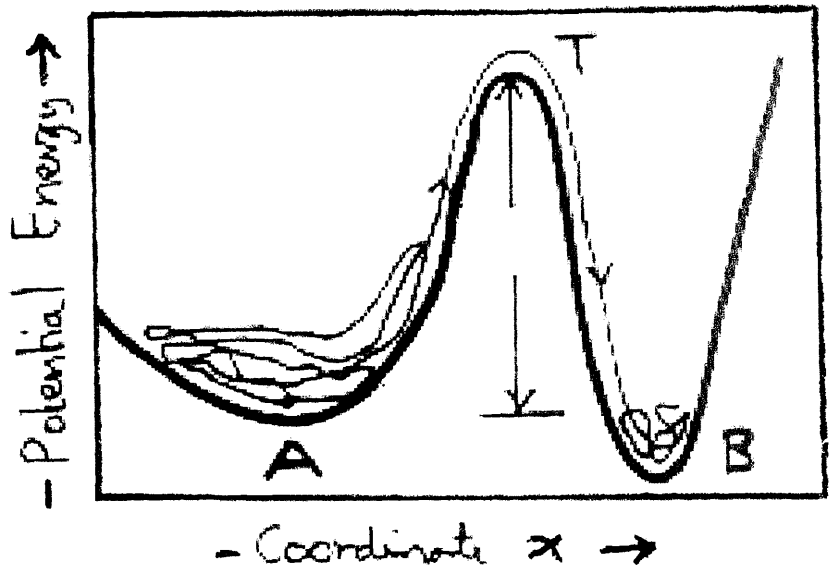
$$\text{normalization} = Q = \sum_i P_i = \sum_i \exp(-U(x_i)/kT)$$

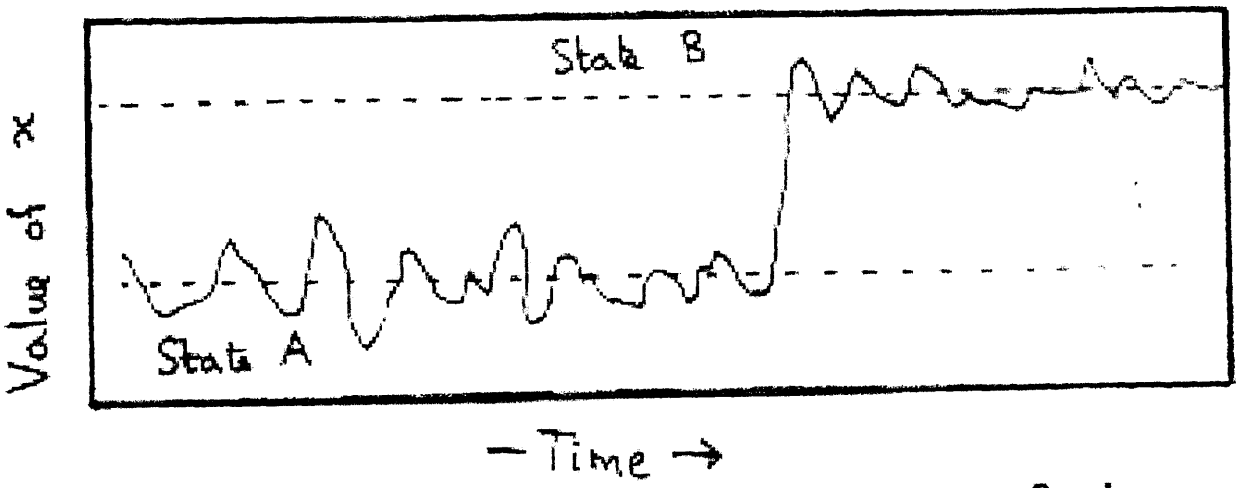
which is called the "partition function".



CROSSING ENERGY BARRIERS

Basic theory of changes of state, chemical reaction. Look at a time-course





The actual transition from A to B is very quick (few pico seconds). What takes time is the waiting. Theory gives the average wait time in state A as

Boltzmann constant in kcal/mol

$$\tau_{A\to B} = \left(\frac{h}{kT}\right) \exp\left(+ (U_T - U_A)/RT \right)$$

Planck's constant

Boltzmann constant

$\left(\frac{h}{kT}\right) \approx 0.16$ picosecond at $T = 300°K$ $(27°C)$.

(remember as $h\nu_0 = kT$. $\nu_0 = 208\ cm^{-1}$, $\tau_{A\to B} = 0.16$ ps)

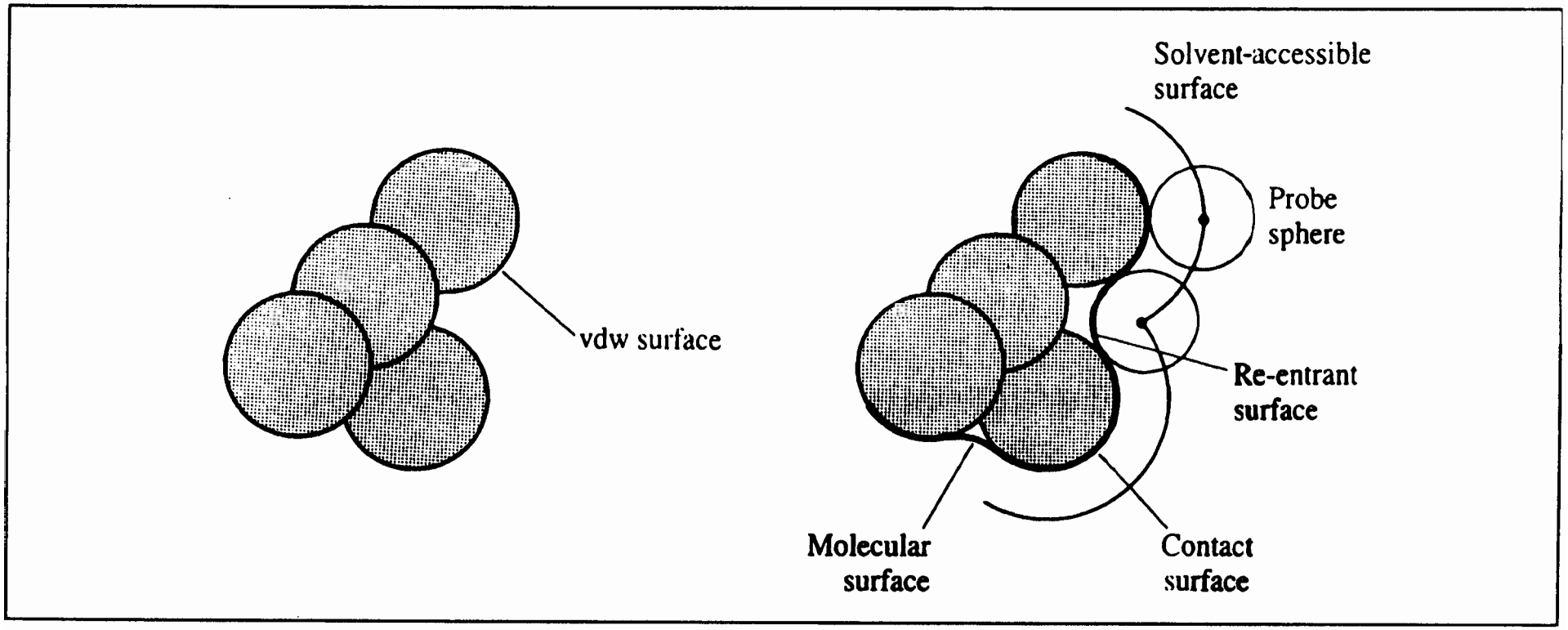


Fig. 1.6: The van der Waals (vdw) surface of a molecule corresponds to the outward-facing surfaces of the van der Waals spheres of the atoms. The molecular surface is generated by rolling a spherical probe (usually of radius 1.4 Å to represent a water molecule) on the van der Waals surface. The molecular surface is constructed from contact and re-entrant surface elements. The centre of the probe traces out the accessible surface.

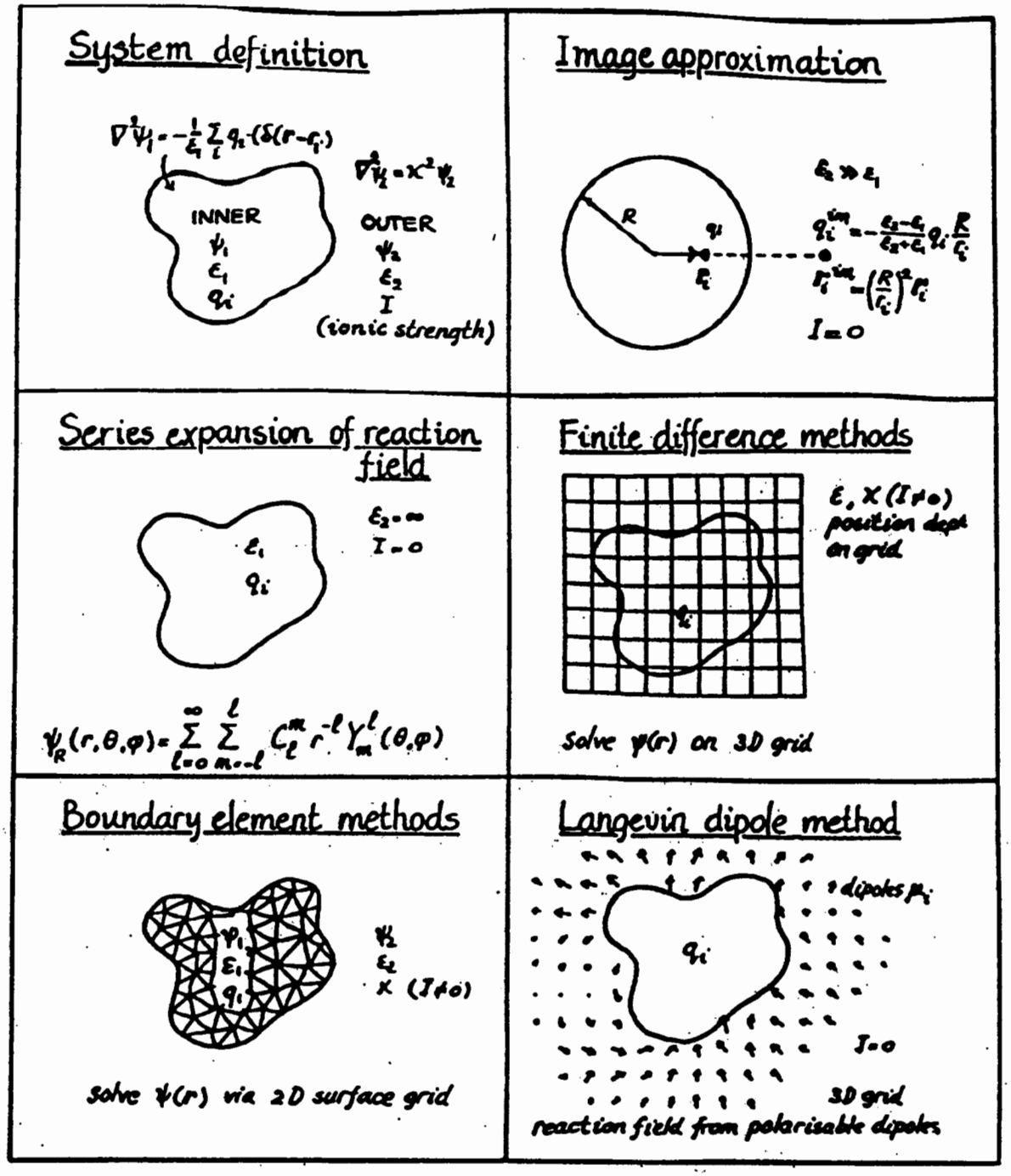


Fig. 8. Non-periodic methods for computing long-range Coulomb forces.