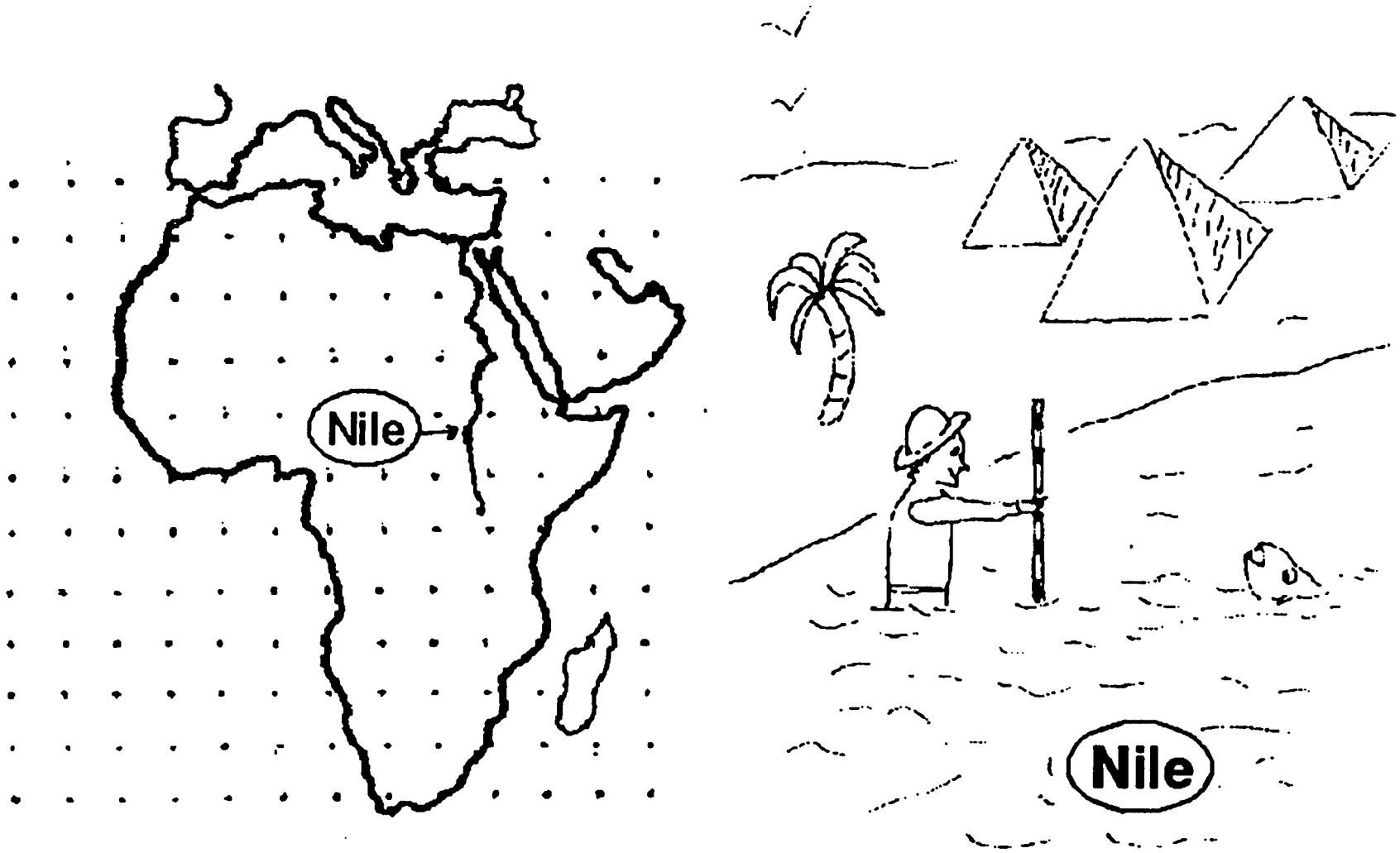


Figure 3.1: Measuring the depth of the Nile: a comparison of conventional quadrature (left), with the Metropolis scheme (right).

How To Do A Monte Carlo Simulation ....

Classical Statistical Mechanics

PROB. OF OCCURENCE — UNIFORM
WEIGHT IN AVERAGE — BOLTZMANN

$$\langle Q \rangle = \sum_J P_J Q_J$$

$$P_J = e^{-\beta E_J} / \sum_J e^{-\beta E_J} \qquad \beta = 1/kT$$

Metropolis Sampling

PROB. OF OCCURENCE — BOLTZMANN
WEIGHT IN AVERAGE — UNIFORM

① Displace $X_J \rightarrow X_{J+1}$

② If $E_{J+1} < E_J$, then keep X_{J+1}.
If $E_{J+1} > E_J$, pick m on $(0,1)$
and if $e^{-\beta(E_{J+1}-E_J)} \geq m$
then keep X_{J+1}, else $X_{J+1} = X_J$.

Monte Carlo Methods:

Edward Teller (1953) J Chem Phys 21 1087

Bill Jorgensen (1978) J Amer Chem Soc 100 7824

---

1. Periodic Boundary Conditions w/ ~100 Molecules
2. Minimum Image or Spherical Cutoff
3. Metropolis Sampling w/ Linear Weighting
4. Pairwise Intramolecular Potentials Used
5. ~500K Configurations Required

Let us now "derive" the Metropolis scheme to determine the transition probability $\pi(o \rightarrow n)$ to go from configuration o to n. It is convenient to start with a thought experiment (actually a thought simulation). We carry out a very large number (say M) Monte Carlo simulations in parallel, where M is much larger than the total number of accessible configurations. We denote the number of points in any configuration o by $m(o)$. We wish that, on average, $m(o)$ is proportional to $\mathcal{N}(o)$. The matrix elements $\pi(o \rightarrow n)$ must satisfy one obvious condition: they do not destroy such an equilibrium distribution once it is reached. This means that, in equilibrium, the average number of accepted trial moves that result in the system leaving state o must be exactly equal to the number of accepted trial moves from all other states n to state o. It is convenient to impose a much stronger condition; namely, that in equilibrium the average number of accepted moves from o to any other state n is exactly canceled by the number of reverse moves. This detailed balance condition implies the following:

$$\mathcal{N}(o)\pi(o \rightarrow n) = \mathcal{N}(n)\pi(n \rightarrow o). \tag{3.1.13}$$

Many possible forms of the transition matrix $\pi(o \rightarrow n)$ satisfy equation (3.1.13). Let us look how $\pi(o \rightarrow n)$ is constructed in practice. We recall that a Monte Carlo move consists of two stages. First, we perform a trial move from state o to state n. We denote the transition matrix that determines the probability to perform a trial move from i to j by $\alpha(o \rightarrow n)$; where α is usually referred to as the underlying matrix of Markov chain [43]. The next stage is the decision to either accept or reject this trial move. Let us denote the probability of accepting a trial move from o to n by $\mathrm{acc}(o \rightarrow n)$. Clearly,

$$\pi(o \rightarrow n) = \alpha(o \rightarrow n) \times \mathrm{acc}(o \rightarrow n). \tag{3.1.14}$$

In the original Metropolis scheme, α is chosen to be a symmetric matrix ($\mathrm{acc}(o \rightarrow n) = \mathrm{acc}(n \rightarrow o)$). However, in later sections we shall see several examples where α is *not* symmetric. If α is symmetric, we can rewrite equation (3.1.13) in terms of the $\mathrm{acc}(o \rightarrow n)$:

$$\mathcal{N}(o) \times \mathrm{acc}(o \rightarrow n) = \mathcal{N}(n) \times \mathrm{acc}(n \rightarrow o). \tag{3.1.15}$$

From equation (3.1.15) follows

$$\frac{\mathrm{acc}(o \rightarrow n)}{\mathrm{acc}(n \rightarrow o)} = \frac{\mathcal{N}(n)}{\mathcal{N}(o)} = \exp\{-\beta[\mathcal{U}(n) - \mathcal{U}(o)]\}. \tag{3.1.16}$$

Again, many choices for $\mathrm{acc}(o \to n)$ satisfy this condition (and the obvious condition that the probability $\mathrm{acc}(o \to n)$ cannot exceed 1). The choice of Metropolis *et al.* is

$$\begin{aligned} \mathrm{acc}(o \to n) &= \mathcal{N}(o)/\mathcal{N}(n) \quad \text{if } \mathcal{N}(n) < \mathcal{N}(o) \\ &= 1 \qquad\qquad\quad \text{if } \mathcal{N}(n) \geq \mathcal{N}(o). \end{aligned} \tag{3.1.17}$$

Other choices for $\mathrm{acc}(o \to n)$ are possible (for a discussion, see for instance [19]), but the original choice of Metropolis *et al.* appears to result in a more efficient sampling of configuration space than most other strategies that have been proposed.

In summary, then, in the Metropolis scheme, the transition probability for going from state o to state n is given by

$$\begin{aligned} \pi(o \to n) &= \alpha(o \to n) & \mathcal{N}(n) \geq \mathcal{N}(o) \\ &= \alpha(o \to n)[\mathcal{N}(n)/\mathcal{N}(o)] & \mathcal{N}(n) < \mathcal{N}(o) \\ \pi(o \to o) &= 1 - \textstyle\sum_{n \neq o} \pi(o \to n). \end{aligned} \tag{3.1.18}$$

Note that we still have not specified the matrix α, except for the fact that it must be symmetric. This reflects considerable freedom in the choice of our trial moves. We will come back to this point in subsequent sections.

One thing that we have not yet explained is how to decide whether a trial move is to be accepted or rejected. The usual procedure is as follows. Suppose that we have generated a trial move from state o to state n, with $\mathcal{U}(n) > \mathcal{U}(o)$. According to equation (3.1.16) this trial move should be accepted with a probability

$$\mathrm{acc}(o \to n) = \exp\{-\beta[\mathcal{U}(n) - \mathcal{U}(o)]\} < 1.$$

In order to decide whether to accept or reject the trial move, we generate a random number, denoted by Ranf, from a uniform distribution in the interval $[0, 1]$. Clearly, the probability that Ranf is less than $\mathrm{acc}(o \to n)$ is equal to $\mathrm{acc}(o \to n)$. We now accept the trial move if $\mathrm{Ranf} < \mathrm{acc}(o \to n)$ and reject it otherwise. This rule guarantees that the probability to accept a trial move from o to n is indeed equal to $\mathrm{acc}(o \to n)$. Obviously, it is very important that our random number generator does indeed generate numbers uniformly in the interval $[0, 1]$. Otherwise the Monte Carlo sampling will be biased.

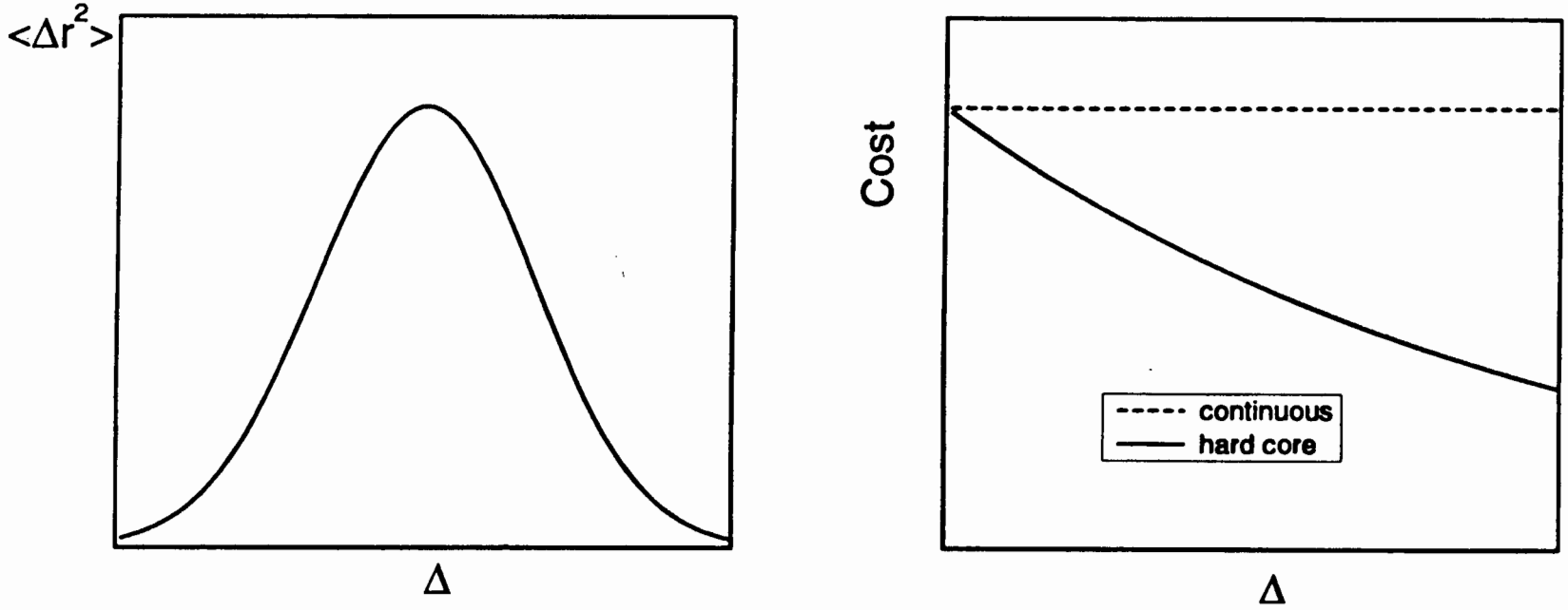


Figure 3.4: (left) Typical dependence of the mean-square displacement of a particle on the average size Δ of the trial move. (right) Typical dependence of the computational cost of a trial move on the step-size Δ. For continuous potentials, the cost is constant, while for hard-core potentials it decreases rapidly with the size of the trial move.

The Traveling Salesman

1. *Configuration.* The cities are numbered $i = 1 \ldots N$ and each has coordinates (x_i, y_i). A configuration is a permutation of the number $1 \ldots N$, interpreted as the order in which the cities are visited.

2. *Rearrangements.* An efficient set of moves has been suggested by Lin [6]. The moves consist of two types: (a) A section of path is removed and then replaced with the same cities running in the opposite order; or (b) a section of path is removed and then replaced in between two cities on another, randomly chosen, part of the path.

3. *Objective Function.* In the simplest form of the problem, E is taken just as the total length of journey,

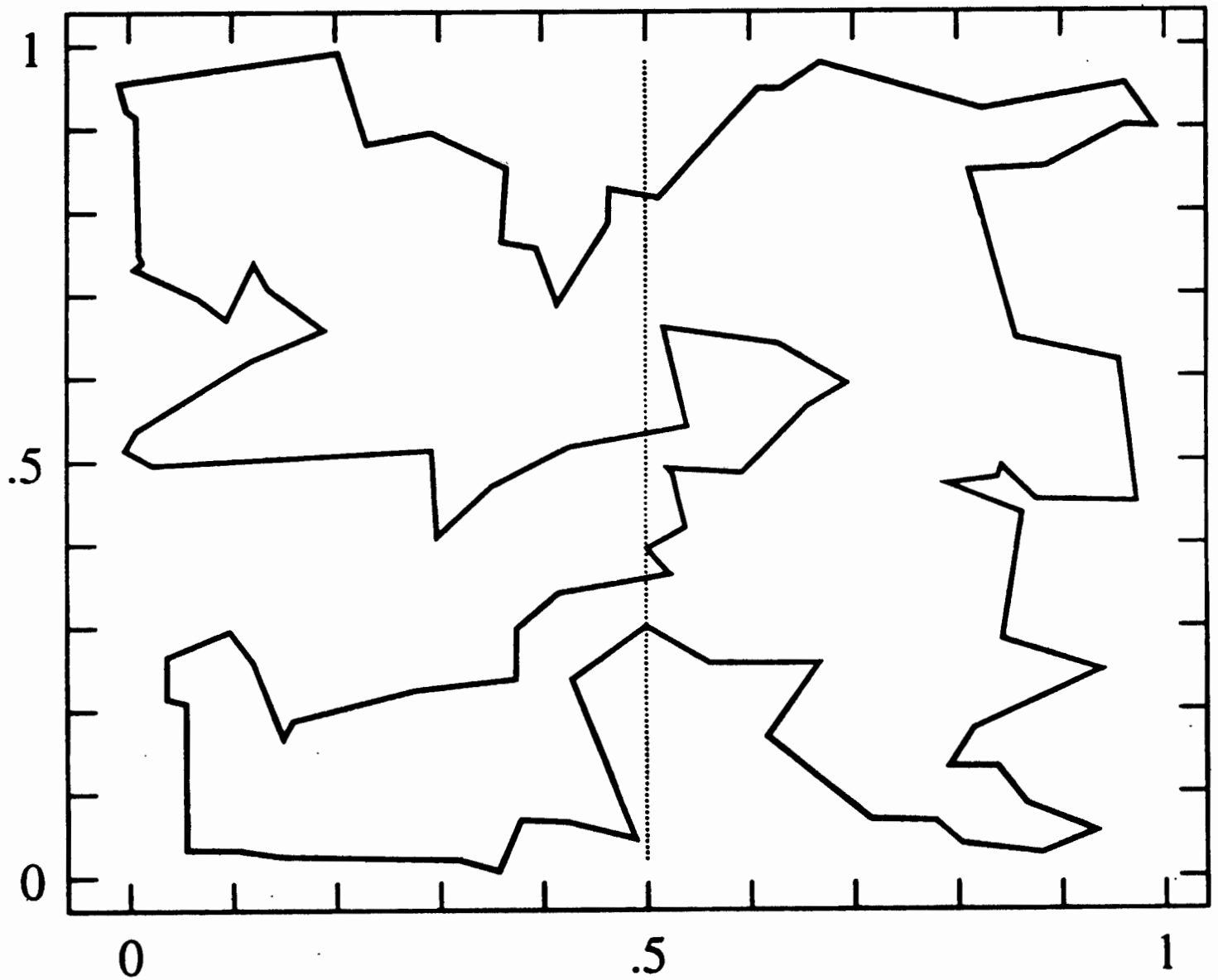
$$E = L \equiv \sum_{i=1}^{N} \sqrt{(x_i - x_{i+1})^2 + (y_i - y_{i+1})^2} \tag{10.9.2}$$

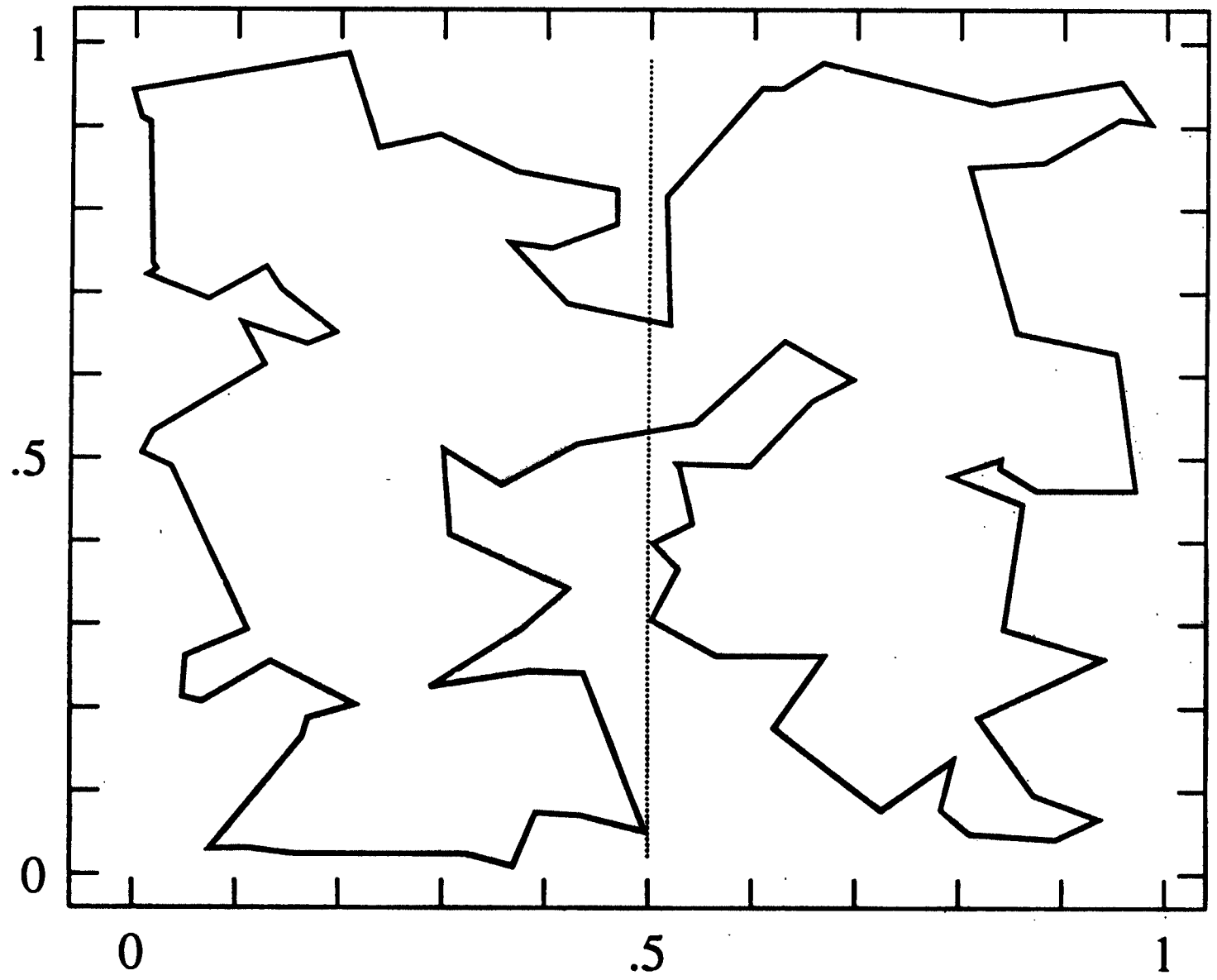
with the convention that point $N + 1$ is identified with point 1. To illustrate the flexibility of the method, however, we can add the following additional wrinkle: Suppose that the salesman has an irrational fear of flying over the Mississippi River. In that case, we would assign each city a parameter μ_i, equal to $+1$ if it is east of the Mississippi, -1 if it is west, and take the objective function to be

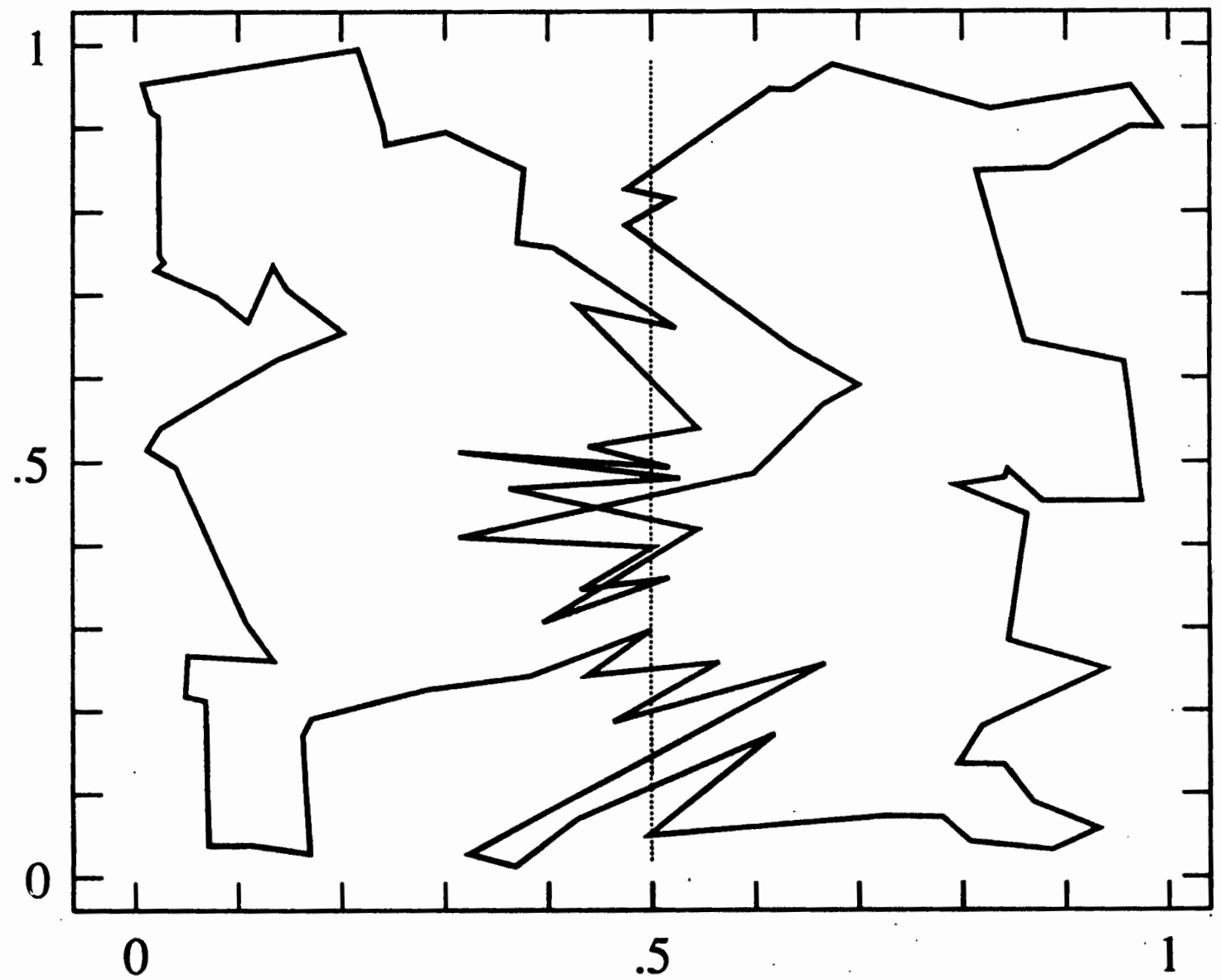
$$E = \sum_{i=1}^{N} \left[\sqrt{(x_i - x_{i+1})^2 + (y_i - y_{i+1})^2} + \lambda(\mu_i - \mu_{i+1})^2 \right] \tag{10.9.3}$$

A penalty 4λ is thereby assigned to any river crossing. The algorithm now finds the shortest path that avoids crossings. The relative importance that it assigns to length of path versus river crossings is determined by our choice of λ. Figure 10.9.1 shows the results obtained. Clearly, this technique can be generalized to include many conflicting goals in the minimization.

4. *Annealing schedule.* This requires experimentation. We first generate some random rearrangements, and use them to determine the range of values of ΔE that will be encountered from move to move. Choosing a starting value for the parameter T which is considerably larger than the largest ΔE normally encountered, we proceed downward in multiplicative steps each amounting to a 10 percent decrease in T. We hold each new value of T constant for, say, $100N$ reconfigurations, or for $10N$ successful reconfigurations, whichever comes first. When efforts to reduce E further become sufficiently discouraging, we stop.


1
.5
0
0
.5
1


1
.5
0
0
.5
1


1
.5
0
0
.5
1

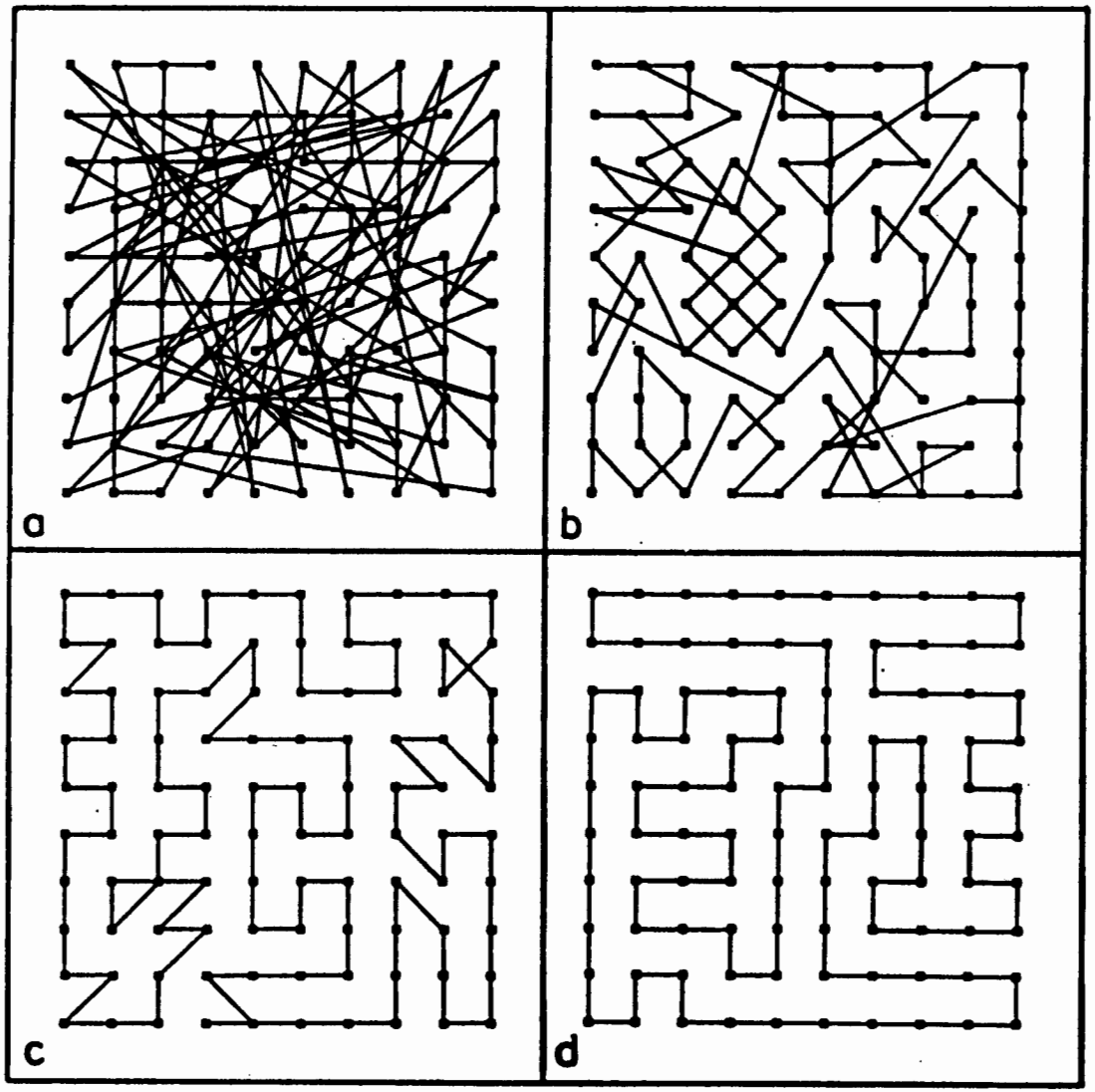


Figure 4.1: Configurations of a 100-city travelling salesman problem obtained after 0, 68, 92 and 123 steps of the algorithm. The initial tour looks very chaotic (a) (high entropy). Gradually, the tour becomes less chaotic, (b) and (c) (the entropy decreases). The final tour (d) shows a highly regular pattern (minimum cost and entropy)

Table II Simulated Annealing[a] of Polyalanines: AcNH-$(Ala)_n$-$CONHCH_3$

n	No. Dihedrals	CPU Time (10 Runs)	Energy (Kj/mole)[b] (No. of Runs)	No. of 1–13 H Bonds[c] (Helical Residues)
2	4	0:24:12	−103.11(5)	0
			− 97.31(5)	
3	6	0:39:16	−151.83(3)	1
			−145.65(3)	
			−140.54(2)	
			−137.78(1)	
			−108.70(1)	
4	8	0:43:28	−210.83(7)	2
			−199.03(1)	2
			−194.81(1)	1
			−186.81(1)	
5	10	0:59:57	−269.48(3)	3
			−256.42(1)	1
			−246.34(1)	
			−236.47(1)	1
			−227.35(2)	
			−219.74(1)	
			−216.92(1)	
6	12	2:34:04	−332.02(5)	4
			−311.85(3)	3
			−294.50(1)	2
			−289.49(1)	1
7	14	1:39:47	−394.95(5)	5
			−374.95(1)	4
			−366.90(1)	3
			−364.96(2)	3
			−336.08(1)	
8	16	2:03:22	−458.42(4)	6
			−437.83(1)	5
			−411.40(1)	4
			−408.89(1)	4
			−409.17(1)	3
			−404.39(1)	3
			−374.28(1)	2
9	18	2:20:18	−521.72(3)	7
			−500.82(1)	6
			−439.54(1)	1
			−436.86(1)	
			−433.05(1)	
			−419.47(1)	1
			−416.88(1)	1
			−395.67(1)	
10	20	2:25:06	−585.63(7)	8
			−544.04(1)	5
			−533.41(1)	4
			−502.33(1)	1

[a]The number of steps in the random walk was 250 at each temperature except for Ala_8 where 500 steps were used.

[b]The lowest energy conformation for cases n = 3–20 was 100% α-helix confirmed by ϕ/ψ angles the characteristic hydrogen-bonding pattern (1–13).

[c]The C=O of residue i is H-bonded to the N—H of residue $i + 4$. A complete α-helix shows $n - 2$ hydrogen bonds. All helices are right-handed.

Table III Simulated Annealing of Larger Polyalanines[a] AcNH-$(Ala)_n$-$CONHCH_3$

n	No. Dihedrals	CPU Time (per Run)	Energy (Kjoul/mol)	No. of 1–13 H Bonds[c] (Helical Residues)
20	40	2:24:00	−1224.07[b]	18
			−1187.75	14
			−1082.72	10
			−1020.12	10
40	80	4:56:06	−2220.05[d]	26
			−1966.49	26
80	160	26:10:30	−4713.81	72
			−4501.94	54

[a]Starting geometry for each run was a sheet with 1 or 2 bends to allow it to fit on the screen.
[b]One-hundred percent α-helix.
[c]The C=O of residue i is H-bonded to the N—H of residue $i + 4$. A complete α-helix shows $n - 2$ hydrogen bonds. All helices are right-handed.
[d]Comparison Ala_{40} α-helix has amber $E = -2475.25$ Kjoul/mol.

Umbrella sampling

Umbrella sampling attempts to overcome the sampling problem by modifying the potential function so that the unfavourable states are sampled sufficiently. The method can be used with both Monte Carlo and molecular dynamics simulations. The modification of the potential function can be written as a perturbation:

$$\mathcal{V}'(\mathbf{r}^N) = \mathcal{V}(\mathbf{r}^N) + W(\mathbf{r}^N)$$

$W(\mathbf{r}^N)$ is a weighting function, which often takes a quadratic form:

$$W(\mathbf{r}^N) = k_W(\mathbf{r}^N - \mathbf{r}_0^N)^2$$

For configurations that are far from the equilibrium state $\mathbf{r}_0^N$ the weighting function will be large and so a simulation using the modified energy function $\mathcal{V}'(\mathbf{r}^N)$ will be biased away from the configuration $\mathbf{r}_0^N$. The resulting distribution will, of course, be non-Boltzmann. The corresponding Boltzmann averages can be extracted from the non-Boltzmann distribution using a method introduced by Torrie and Valleau [Torrie and Valleau 1977]. The result is:

$$\langle A \rangle = \frac{\langle A(\mathbf{r}^N) \exp[+W(\mathbf{r}^N)/k_B T] \rangle_W}{\langle \exp[+W(\mathbf{r}^N)/k_B T] \rangle_W}$$

The subscript W indicates that the average is based on the probability $P_W(\mathbf{r}^N)$, which in turn is determined by the modified energy function $\mathcal{V}'(\mathbf{r}^N)$.

Step 1: Carry out a conformational search to find the set of low-energy conformers—call these $\mathbf{X}_i$. Evaluate the internal coordinate transformations that interconvert all pairs (i,j) of the conformers in the $\mathbf{X}_i$ list—call these transformations $\mathbf{T}_{ij}$.

Step 2: Pick an initial conformation—call this structure $\mathbf{Y}_0$.

Step 3: Find the conformer on the $\mathbf{X}_i$ list that is closest to $\mathbf{Y}_0$—call this conformer $\mathbf{X}_0$.

Step 4: Randomly choose a conformer from the $\mathbf{X}_i$ list—call this conformer $\mathbf{X}_T$.

Step 5: Apply transformation $\mathbf{T}\mathbf{x}_0\mathbf{x}_T$ to structure $\mathbf{Y}_0$ to generate structure $\mathbf{Y}_1$.

Step 6: Apply small random variations to internal coordinates of $\mathbf{Y}_1$ to generate the new trial structure $\mathbf{Y}_2$.

Step 7: Compare energies of $\mathbf{Y}_0$ and $\mathbf{Y}_2$, accepting $\mathbf{Y}_2$ with a probability defined by Metropolis;[3] $p = \min\{1 , \exp[-(E(\mathbf{Y}_2) - E(\mathbf{Y}_0))/kT]\}$.

Step 8: Define the resulting structure as $\mathbf{Y}_0$ and go back to Step 3.

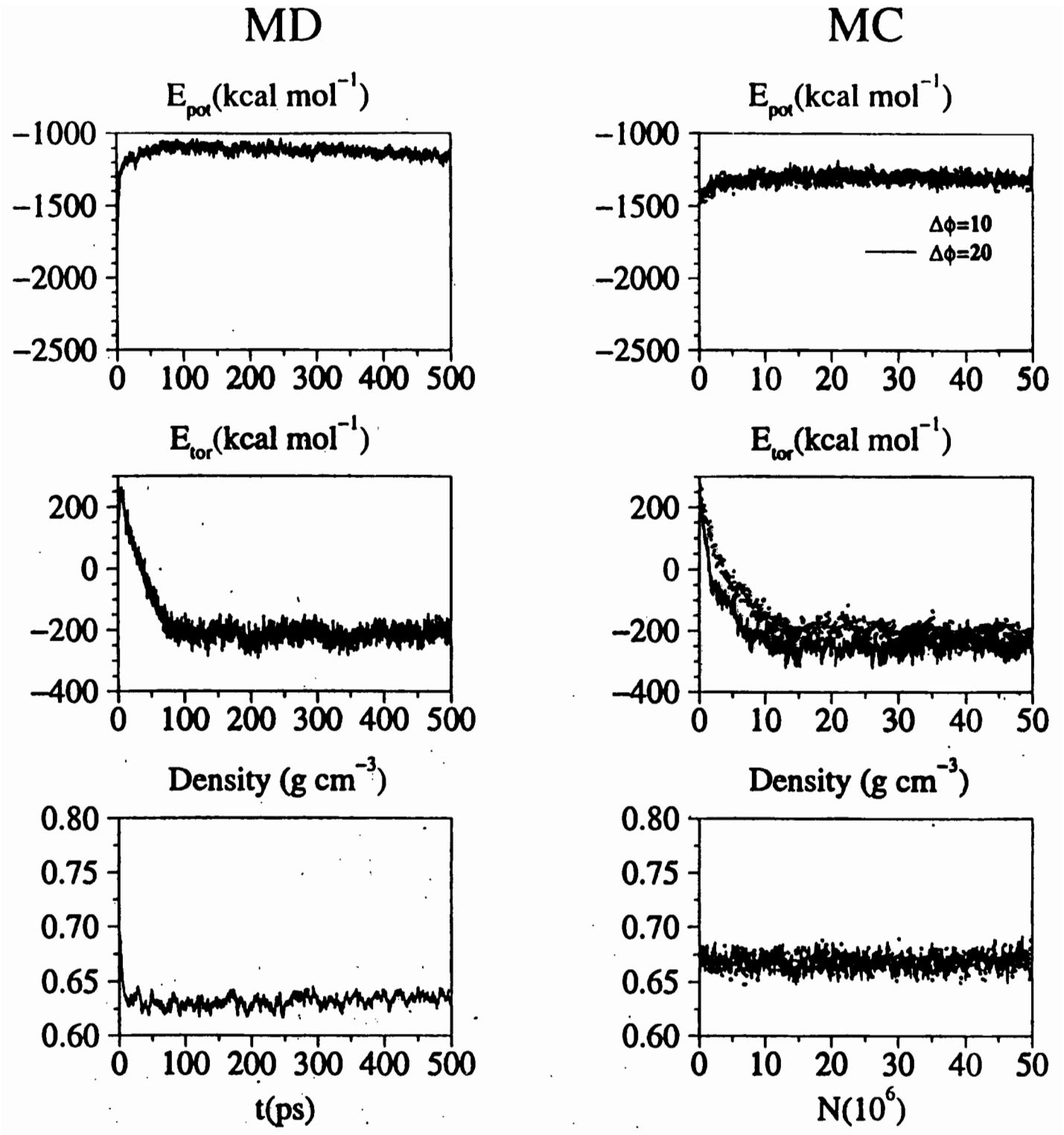


Figure 3. Convergence plots from the MD (left) and MC10 and MC20 (right) simulations for the potential energy, torsional energy, and density.

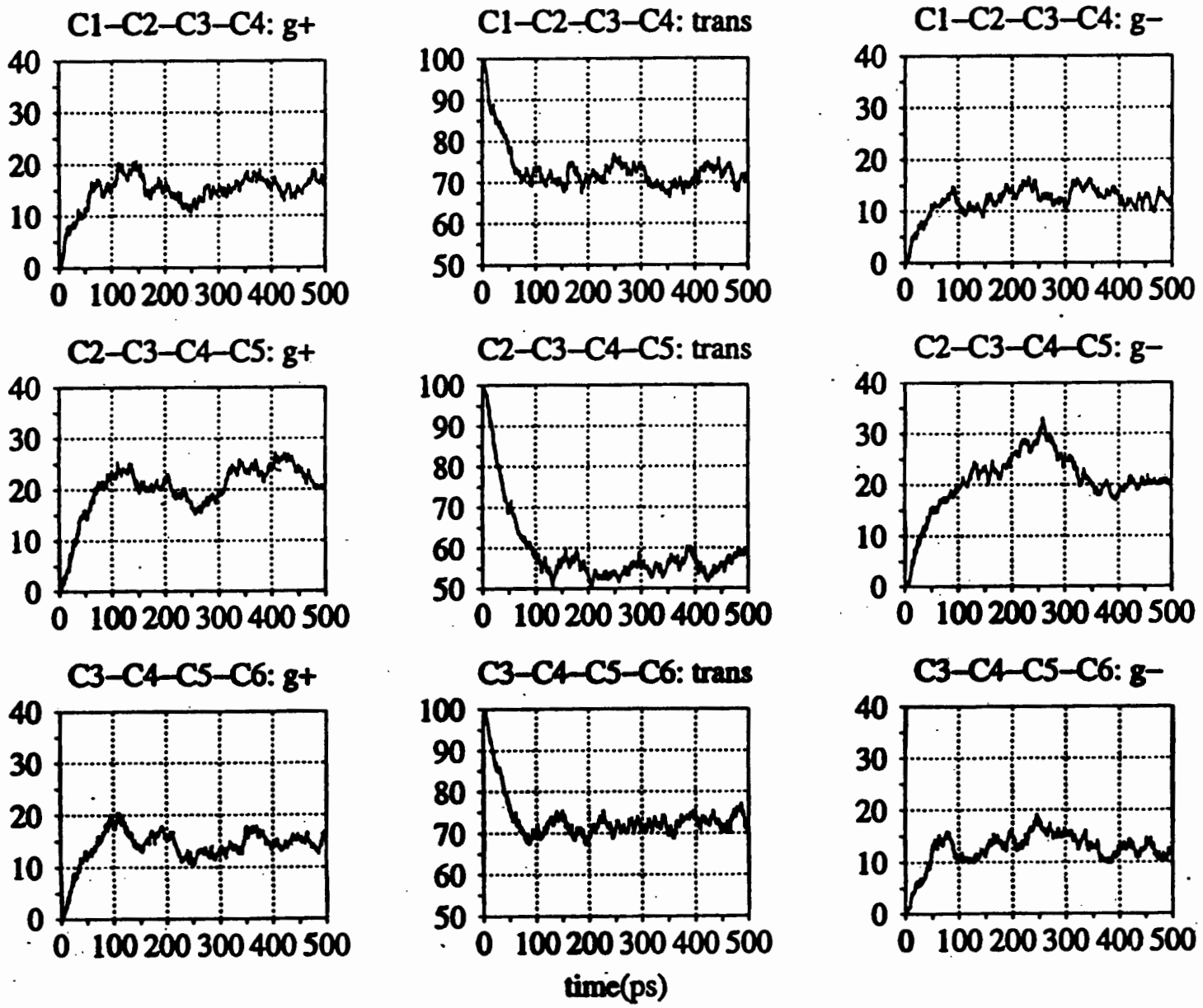


Figure 5. Convergence plots for the *trans* and *gauche* populations in percent for each dihedral angle of hexane from the MD simulation.

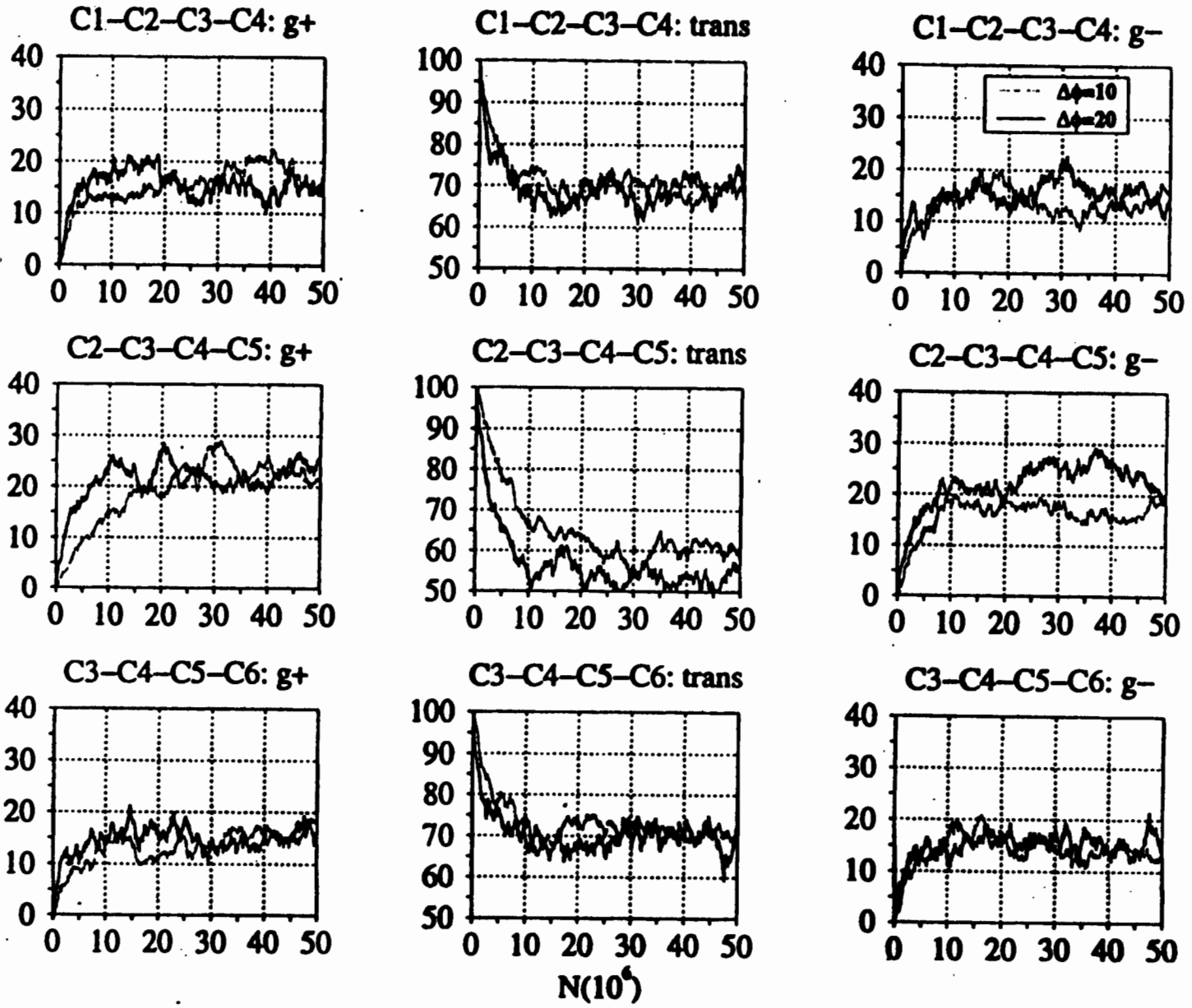


Figure 4. Convergence plots for the *trans* and *gauche* populations in percent for each dihedral angle of hexane from the MC10 and MC20 simulations.

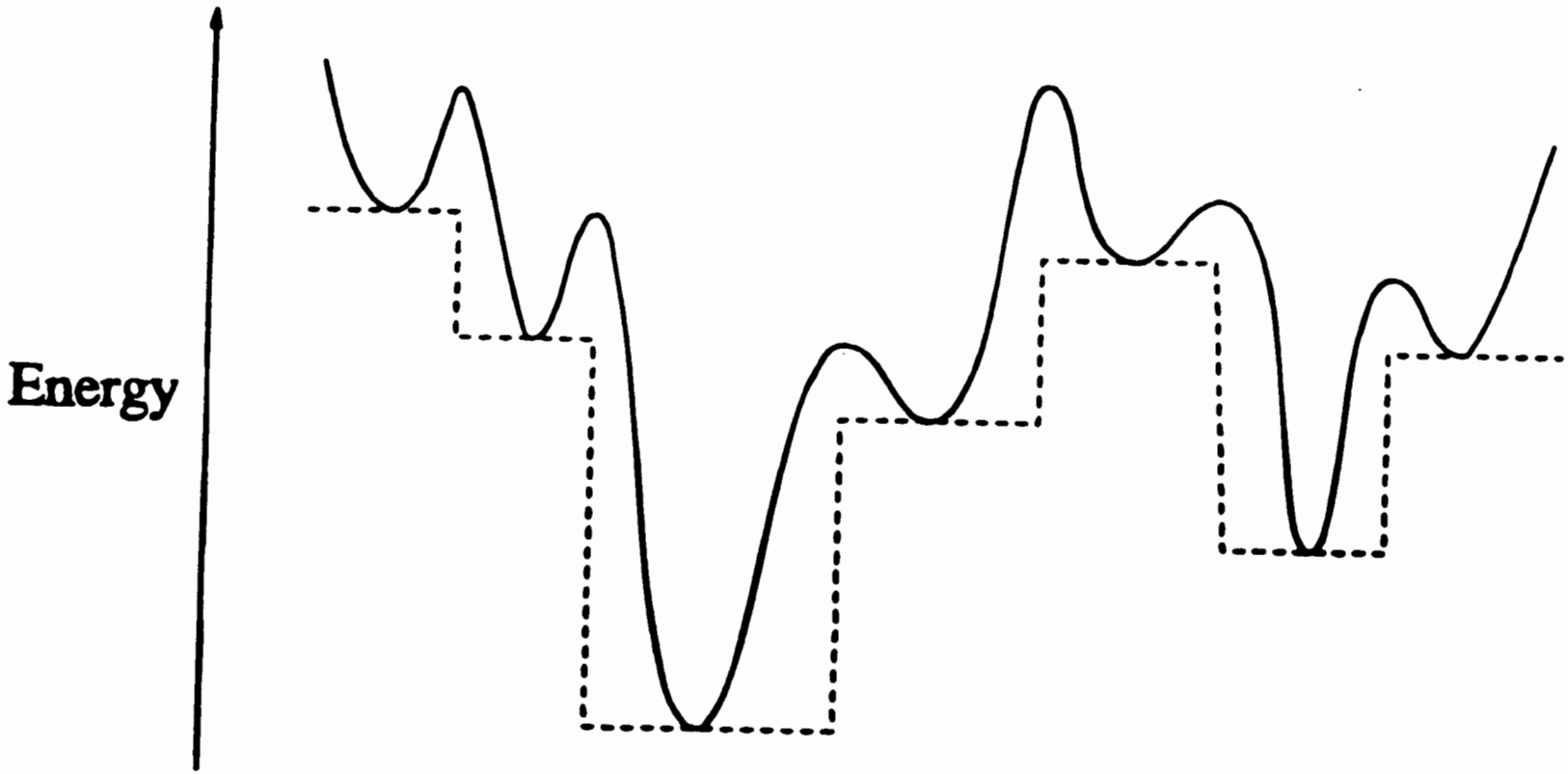


Figure 2. A schematic diagram illustrating the effects of our energy transformation for a one-dimensional example. The solid line is the energy of the original surface and the dashed line is the transformed energy $\tilde{E}$.

TABLE 1: Global Minima of LJ_N for $N \leq 110$

N	point group	energy/ϵ	ref[a]	N	point group	energy/ϵ
2	$D_{\infty h}$	−1.000 000	2	57	C_s	−288.342 625
3	D_{3h}	−3.000 000	2/3	58	C_{3v}	−294.378 148
4	T_d	−6.000 000	2/3	59	C_{2v}	−299.738 070
5	D_{3h}	−9.103 852	2/3	60	C_s	−305.875 476
6	O_h	−12.712 062	2/3	61	C_{2v}	−312.008 896
7	D_{5h}	−16.505 384	2/3	62	C_s	−317.353 901
8	C_s	−19.821 489	2/3	63	C_1	−323.489 734
9	C_{2v}	−24.113 360	2/3	64	C_s	−329.620 147
10	C_{3v}	−28.422 532	2	65	C_2	−334.971 532
11	C_{2v}	−32.765 970	2	66	C_1	−341.110 599
12	C_{5v}	−37.967 600	2	67	C_s	−347.252 007
13	I_h	−44.326 801	2/3	68	C_1	−353.394 542
14	C_{3v}	−47.845 157	2/3	69	C_{5v}	−359.882 566
15	C_{2v}	−52.322 627	2	70	C_{5v}	−366.892 251
16	C_s	−56.815 742	2	71	C_{5v}	−373.349 661
17	C_2	−61.317 995	7	72	C_s	−378.637 253
18	C_{5v}	−66.530 949	2	73	C_s	−384.789 377
19	D_{5h}	−72.659 782	2	74	C_s	−390.908 500
20	C_{2v}	−77.177 043	2	75	D_{5h}	−397.492 331
21	C_{2v}	−81.684 571	2	76	C_s	−402.894 866
22	C_s	−86.809 782	10	77	C_{2v}	−409.083 517
23	D_{3h}	−92.844 472	8	78	C_s	−414.794 401
24	C_s	−97.348 815	9	79	C_{2v}	−421.810 897
25	C_s	−102.372 663	2	80	C_s	−428.083 564
26	T_d	−108.315 616	2	81	C_{2v}	−434.343 643
27	C_{2v}	−112.873 584	10	82	C_1	−440.550 425
28	C_s	−117.822 402	10	83	C_{2v}	−446.924 094
29	D_{3h}	−123.587 371	2	84	C_1	−452.657 214
30	C_{2v}	−128.286 571	10	85	C_{3v}	−459.055 799
31	C_s	−133.586 422	10	86	C_1	−465.384 493
32	C_{2v}	−139.635 524	10	87	C_s	−472.098 165
33	C_s	−144.842 719	10	88	C_s	−479.032 630
34	C_{2v}	−150.044 528	10	89	C_{3v}	−486.053 911
35	C_1	−155.756 643	10	90	C_s	−492.433 908
36	C_s	−161.825 363	10	91	C_s	−498.811 060
37	C_1	−167.033 672	10	92	C_{3v}	−505.185 309
38	O_h	−173.928 427	13/14	93	C_1	−510.877 688
39	C_{5v}	−180.033 185	10	94	C_1	−517.264 131
40	C_s	−185.249 839	10	95	C_1	−523.640 211
41	C_s	−190.536 277	10	96	C_1	−529.879 146
42	C_s	−196.277 534	10	97	C_1	−536.681 383
43	C_s	−202.364 664	10	98	C_s	−543.642 957
44	C_1	−207.688 728	10	99	C_{2v}	−550.666 526
45	C_1	−213.784 862	10	100	C_s	−557.039 820
46	C_{2v}	−220.680 330	10	101	C_{2v}	−563.411 308
47	C_1	−226.012 256	10	102	C_{2v}	−569.363 652
48	C_s	−232.199 529	10	103	C_s	−575.766 131
49	C_{3v}	−239.091 864	10	104	C_{2v}	−582.086 642
50	C_s	−244.549 926	10	105	C_1	−588.266 501
51	C_{2v}	−251.253 964	10	106	C_1	−595.061 072
52	C_{3v}	−258.229 991	10	107	C_s	−602.007 110
53	C_{2v}	−265.203 016	10	108	C_s	−609.033 011
54	C_{5v}	−272.208 631	10	109	C_1	−615.411 166
55	I_h	−279.248 470	4	110	C_s	−621.788 224
56	C_{3v}	−283.643 105	10			

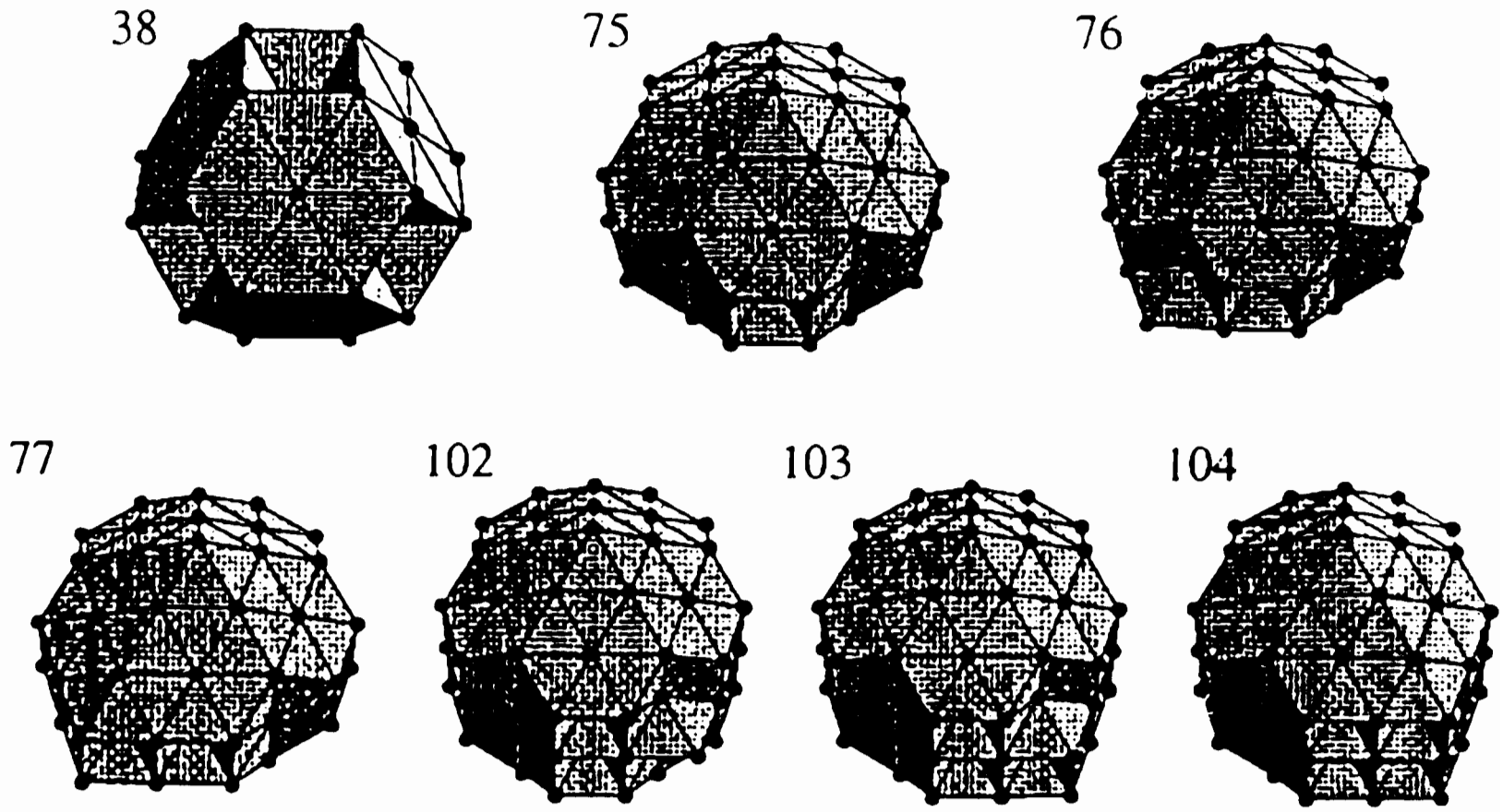


Figure 1. Nonicosahedral Lennard-Jones global minima.

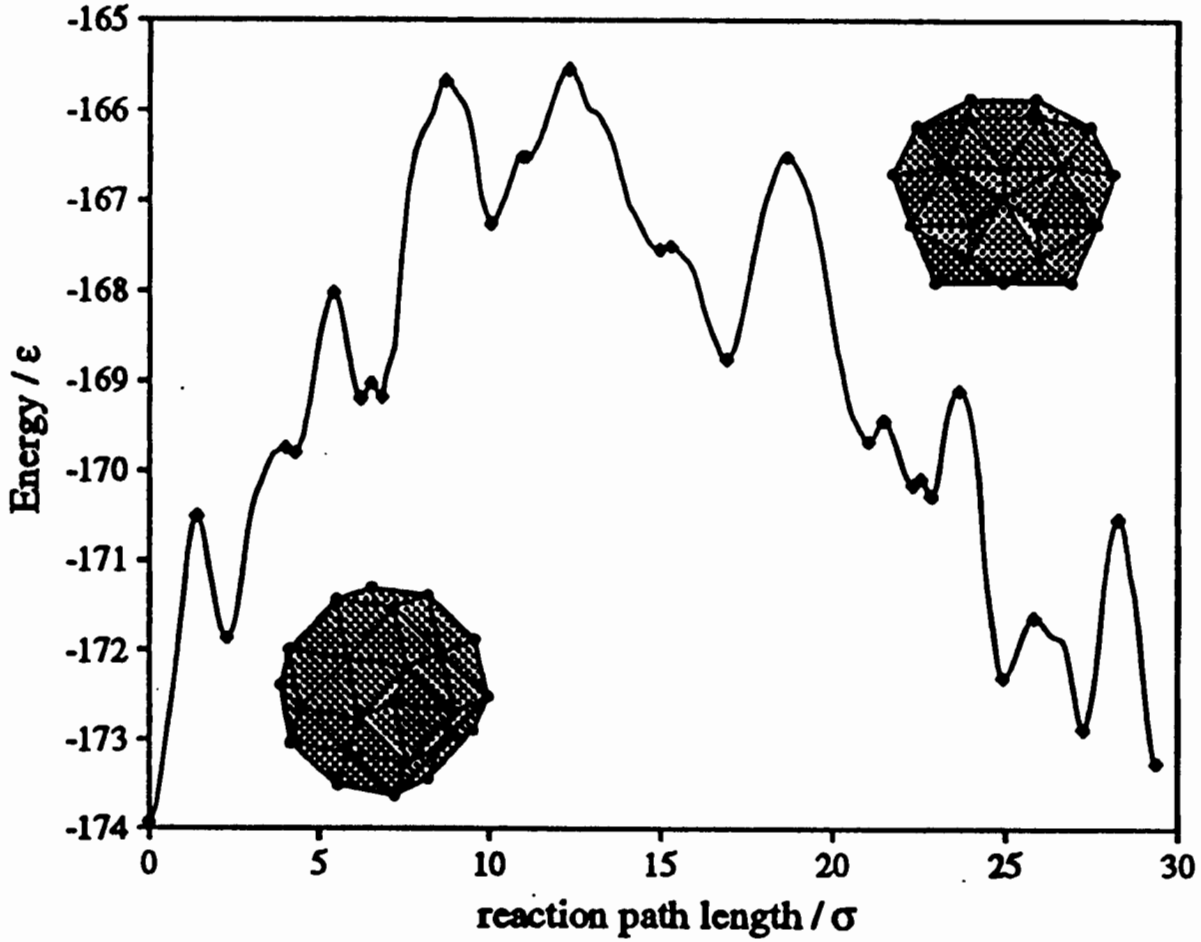


FIG. 1. Energy profile of a pathway between the two lowest energy minima of LJ_{38}, namely the fcc truncated octahedron (bottom left) and a structure based on the Mackay icosahedron with C_{5v} point group symmetry (top right). $2^{1/6}\sigma$ is the equilibrium pair separation of the LJ potential. The method by which this pathway was obtained is described in Ref. [15].

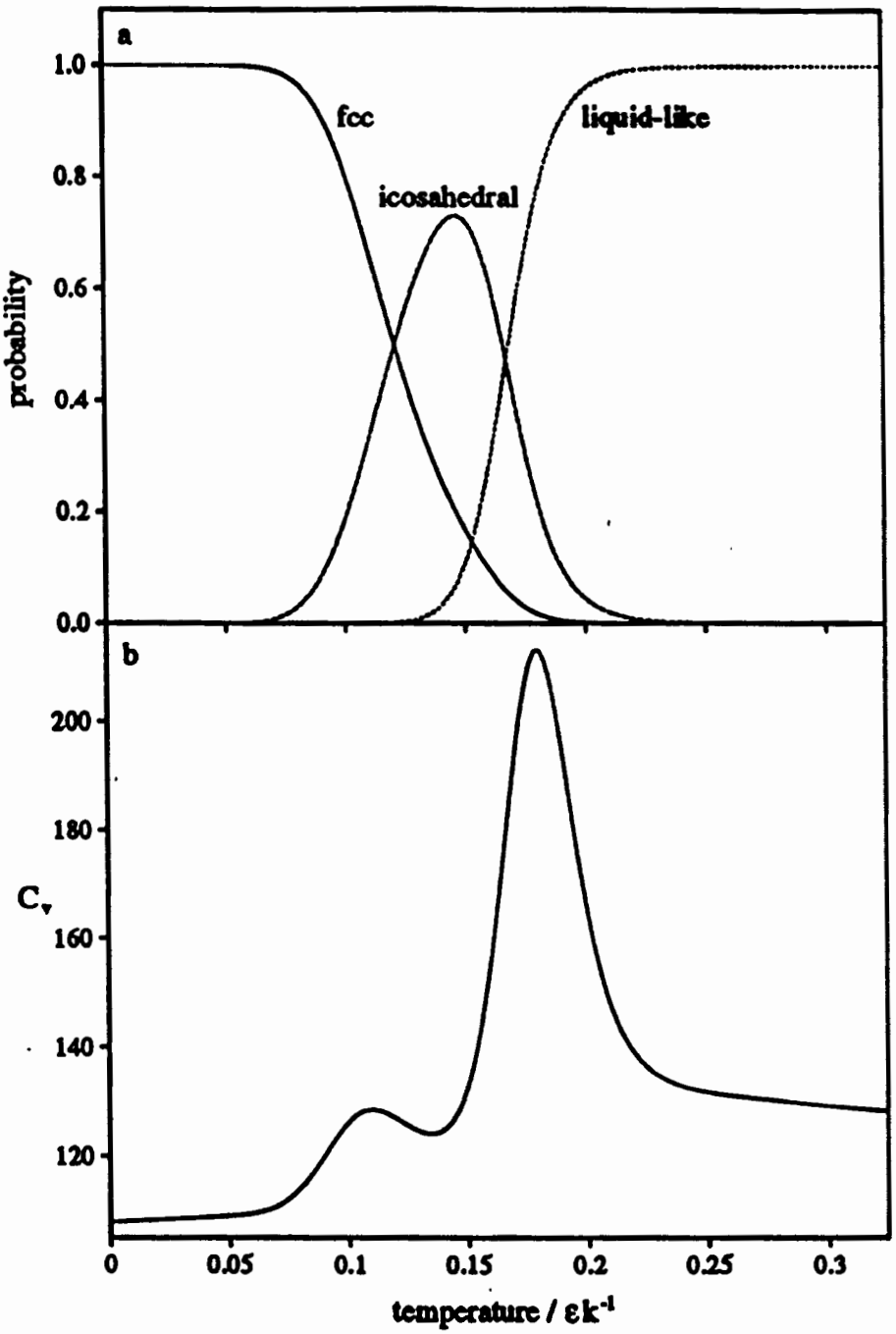


FIG. 2. Equilibrium thermodynamic properties of the untransformed LJ_{38} PES. (a) The probability of the cluster being in the fcc, icosahedral and 'liquid-like' regions of bound configuration space. (b) The heat capacity, C_v. These results were obtained by summing the anharmonic partition functions for a sample of minima appropriately weighted to compensate for the incompleteness of the sample[16]. The liquid-like region of configuration space is defined as those minima with $E > -171.6\,\epsilon$.

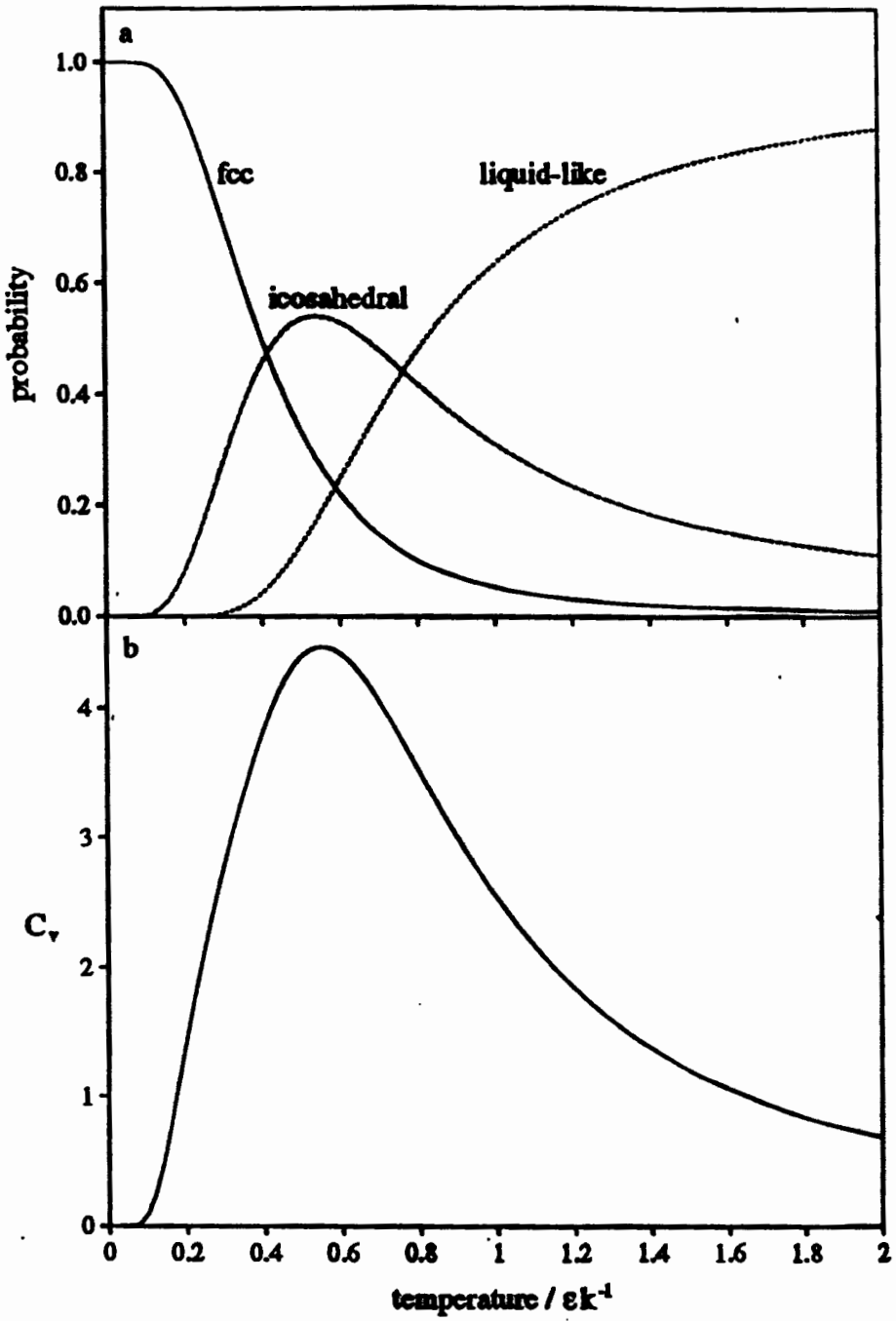


FIG. 3. Equilibrium thermodynamic properties of the transformed LJ_{38} PES. (a) The probability of the cluster being in the fcc, icosahedral and 'liquid-like' regions of bound configuration space. (b) The configurational component of the heat capacity.

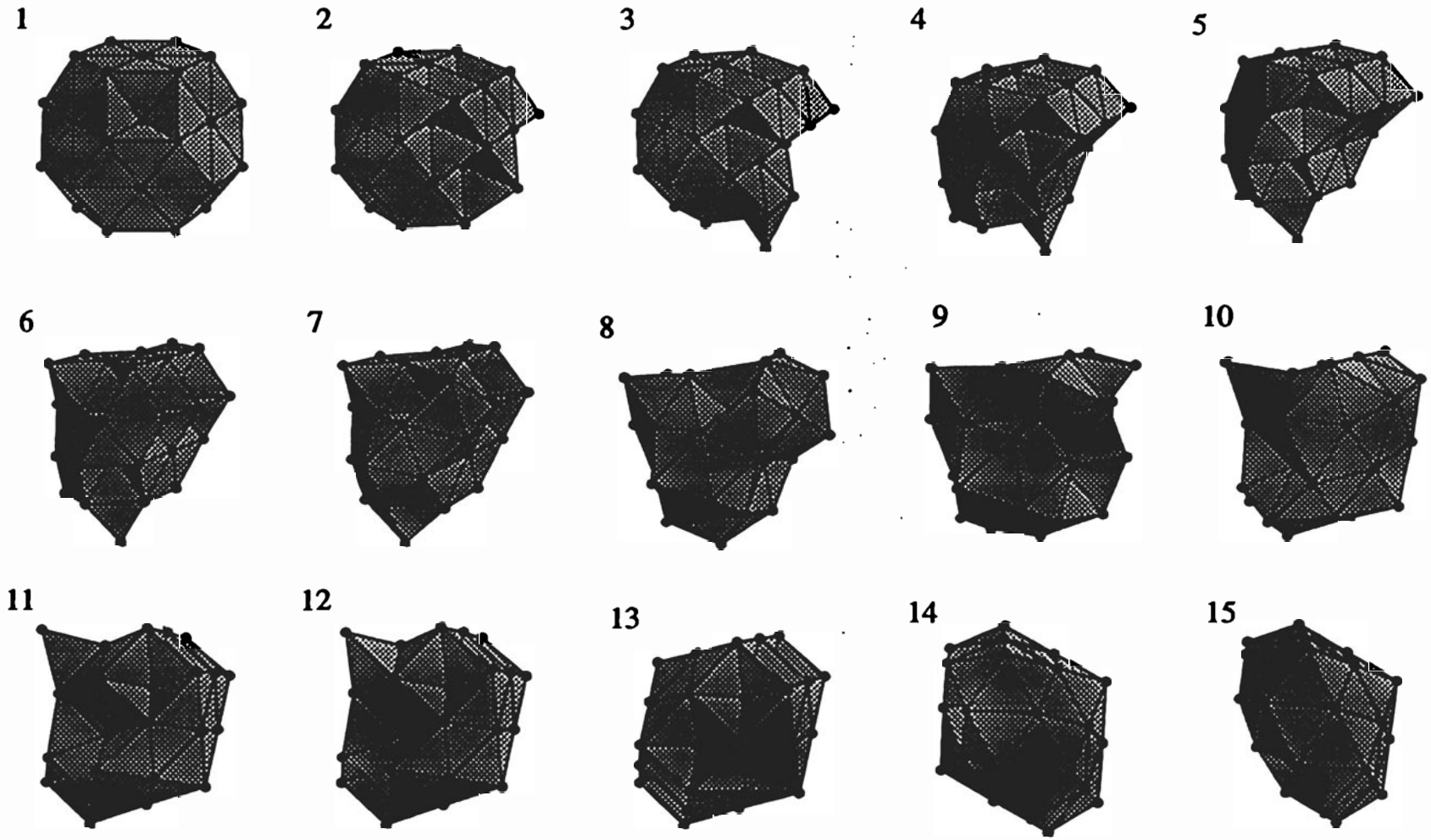


Fig. 2. Minima on the LJ_{38} pathway from the truncated octahedron to the lowest energy icosahedral minimum. The minima are numbered by their position along the pathway

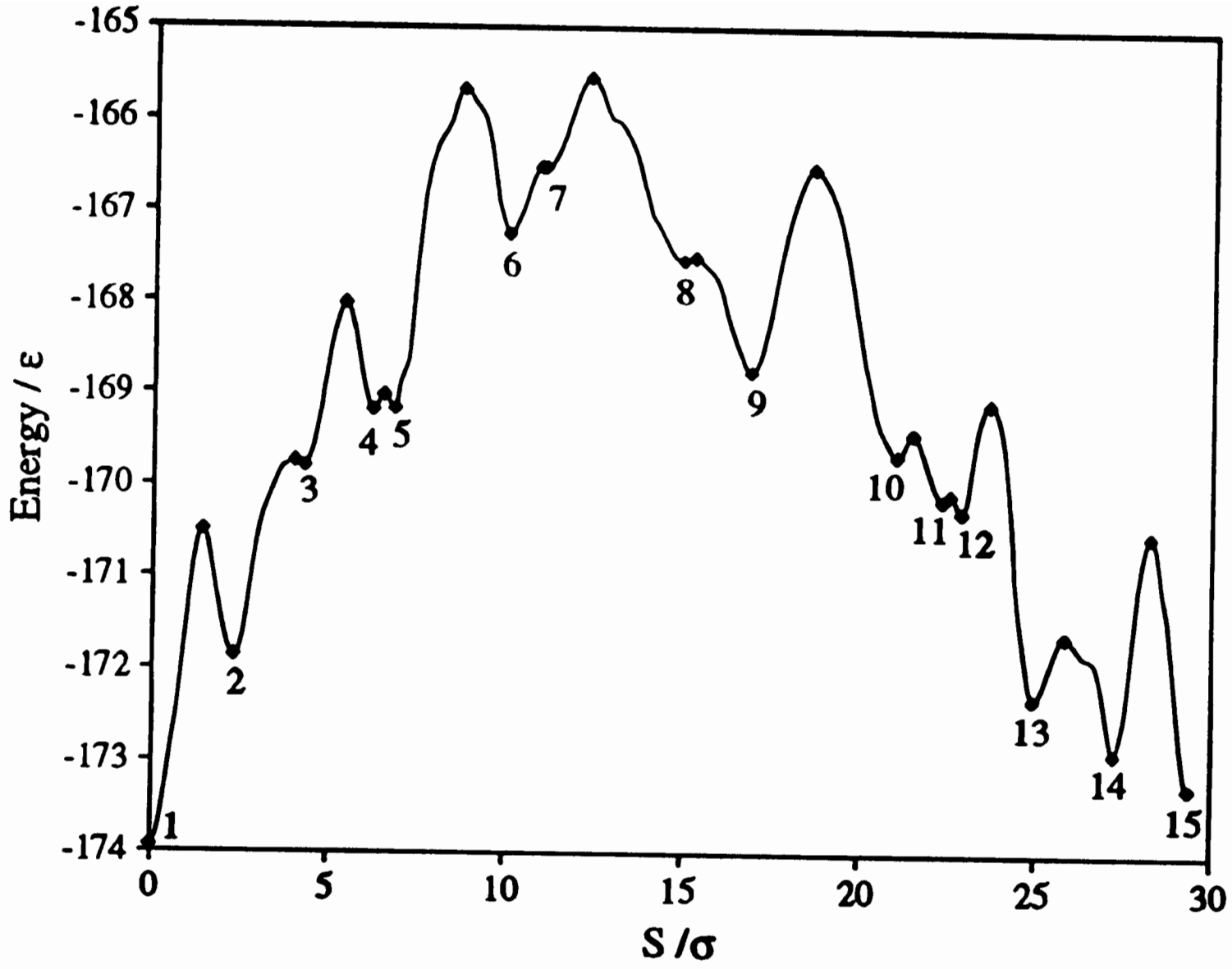


Fig. 3. Reaction profile of a pathway on the LJ_{38} PES from the truncated octahedron to the lowest energy icosahedral minimum. Stationary points are denoted by *diamonds*. The minima are numbered by their position along the pathway