The Unbearable Lightness of Being: An Introduction to The Role of Softness in Biological Molecules

Robert Austin
Visiting Member
Institute for Advanced Study at UST
Professor of Physics, Princeton University

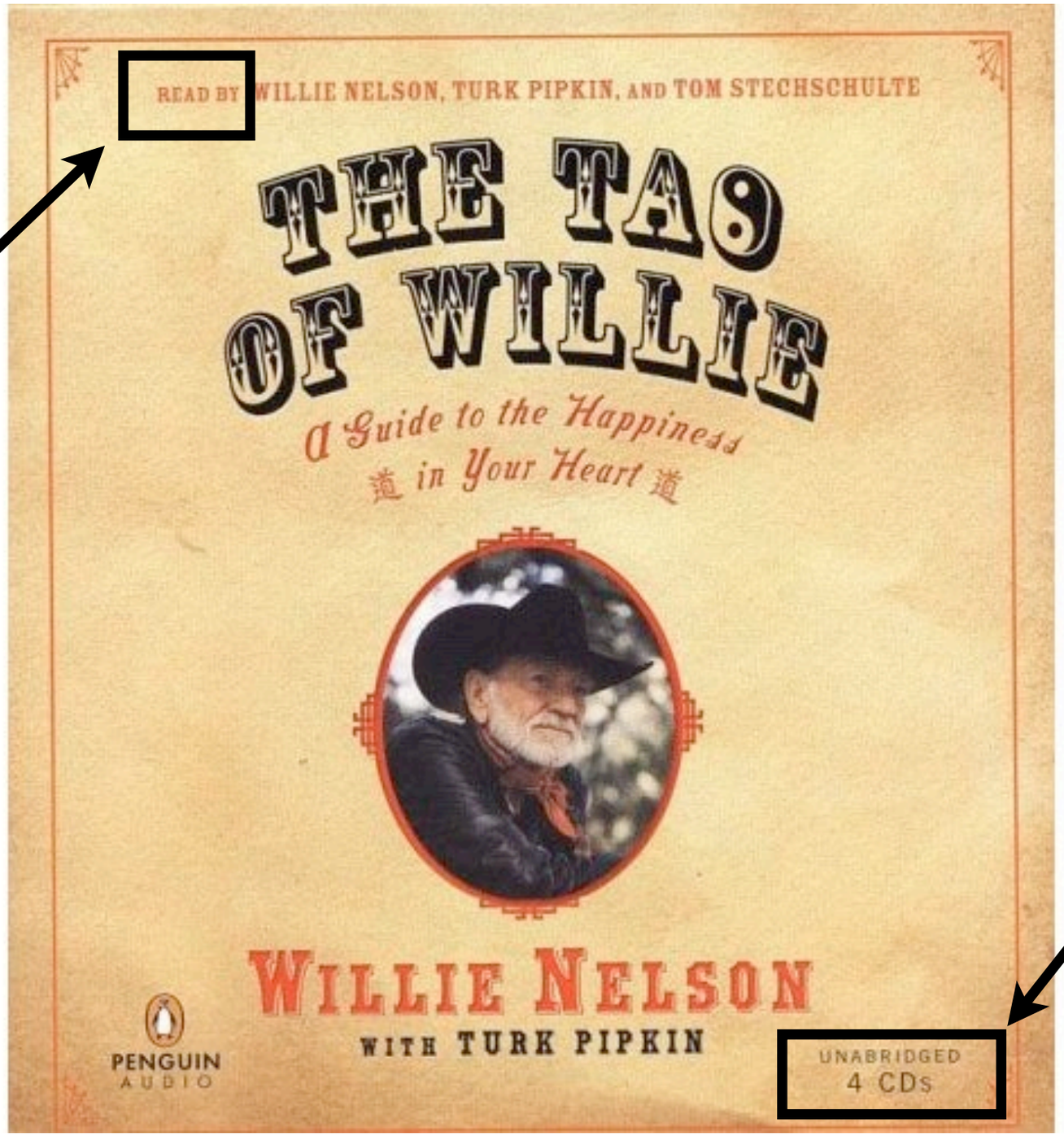
第四十三章

天下之至柔，馳騁天下之至堅。無有入無間，吾是以知無為之有益。不言之教，無為之益，天下希及之。

老子


The softest thing in the world
dashes against and overcomes the
hardest;

PS: Quoting from The Tao is not only irritating to other people, but it is boring, useless advice.



Abstract:

Schrödinger speculated in his epochal 1944 monograph "What is Life?" that biological molecules, and the genetic material in particular, must be some sort of "a-periodic crystal".

We have come a long way since then thanks to the tools of physics, but the idea of "crystals" still holds sway when talking about protein and DNA structure.

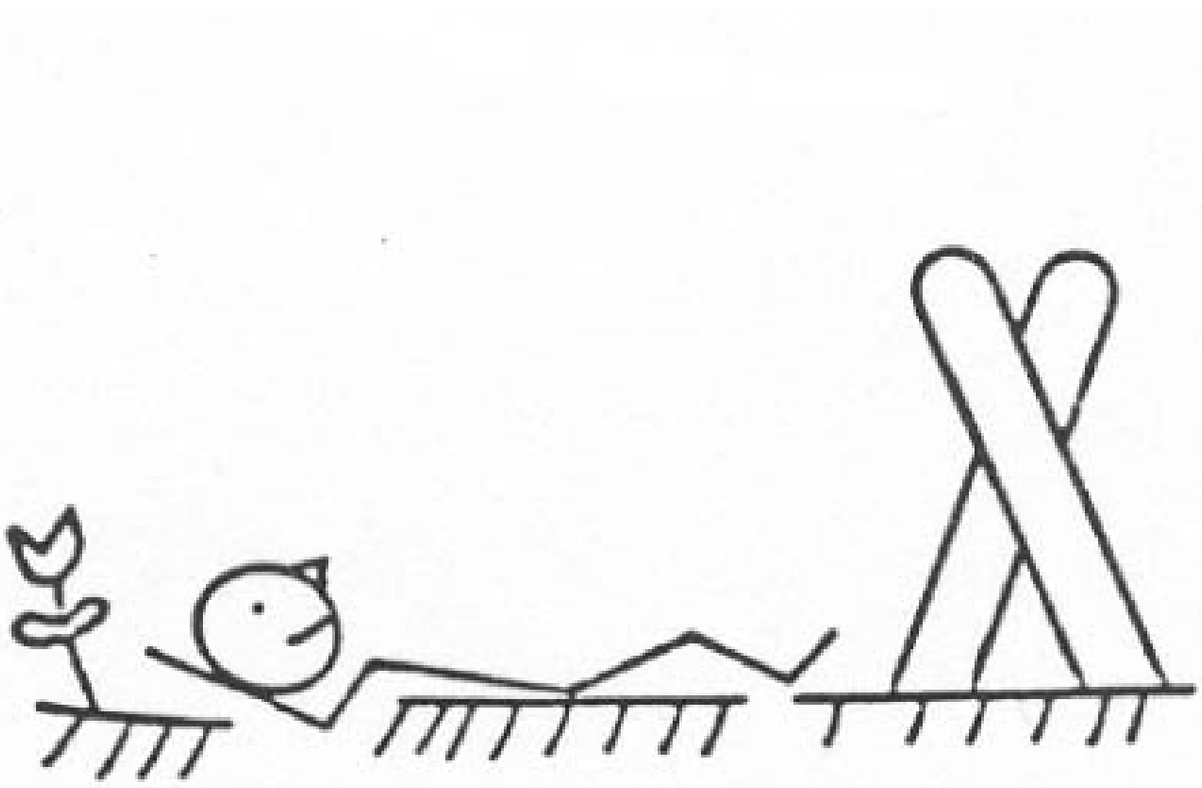
In fact, biological molecules are conformationally flexible: they are soft. This softness is extremely important in understanding how biological molecules function,. I will discuss the role that softness plays in the "being" of biology using work from our lab over the years.

1) The conformational flexibility of proteins and why it matters.

2) The softness of DNA and why it matters.

3) Peering into the future.

1) The conformational flexibility of proteins and why it matters.



Hans Frauenfelder

The Physics of Proteins

Editors: Shirley S. Chan and Winnie S. Chan

With: Robert H. Austin, Uli Nienhaus
Charles G. Schultz and Robert D. Young
March 2009

Springer

Berlin Heidelberg New York
Hong Kong London
Milan Paris Tokyo

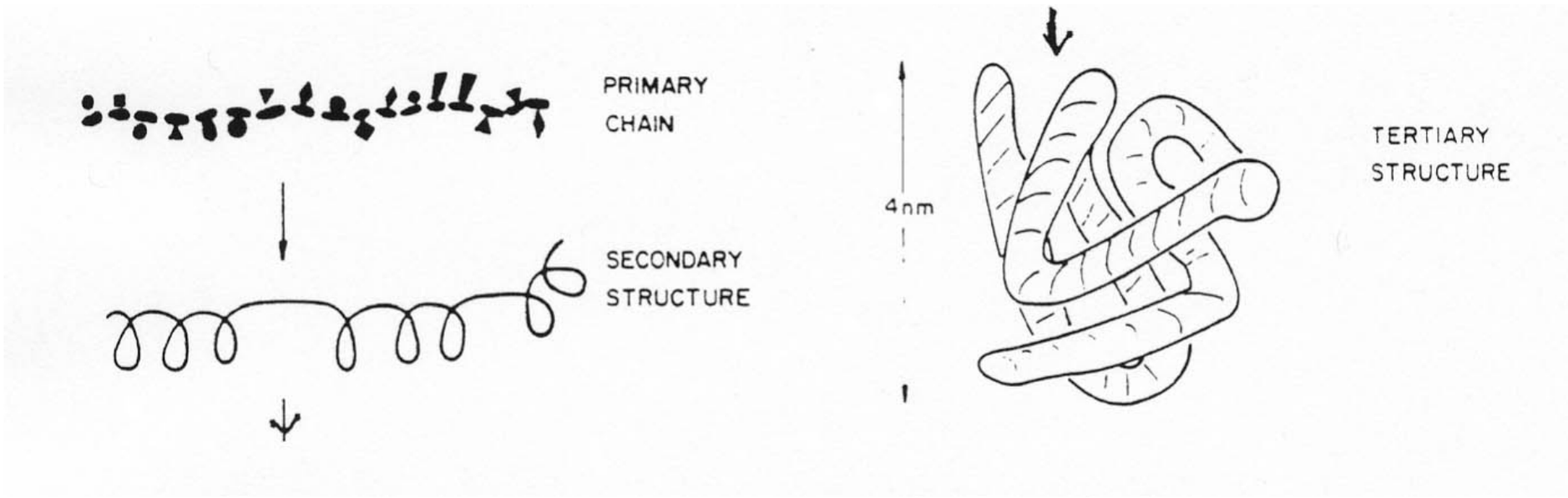


Fig. 4.1. The linear polypeptide chain (primary sequence) folds into the final tertiary structure with a diameter of about 4 nm.

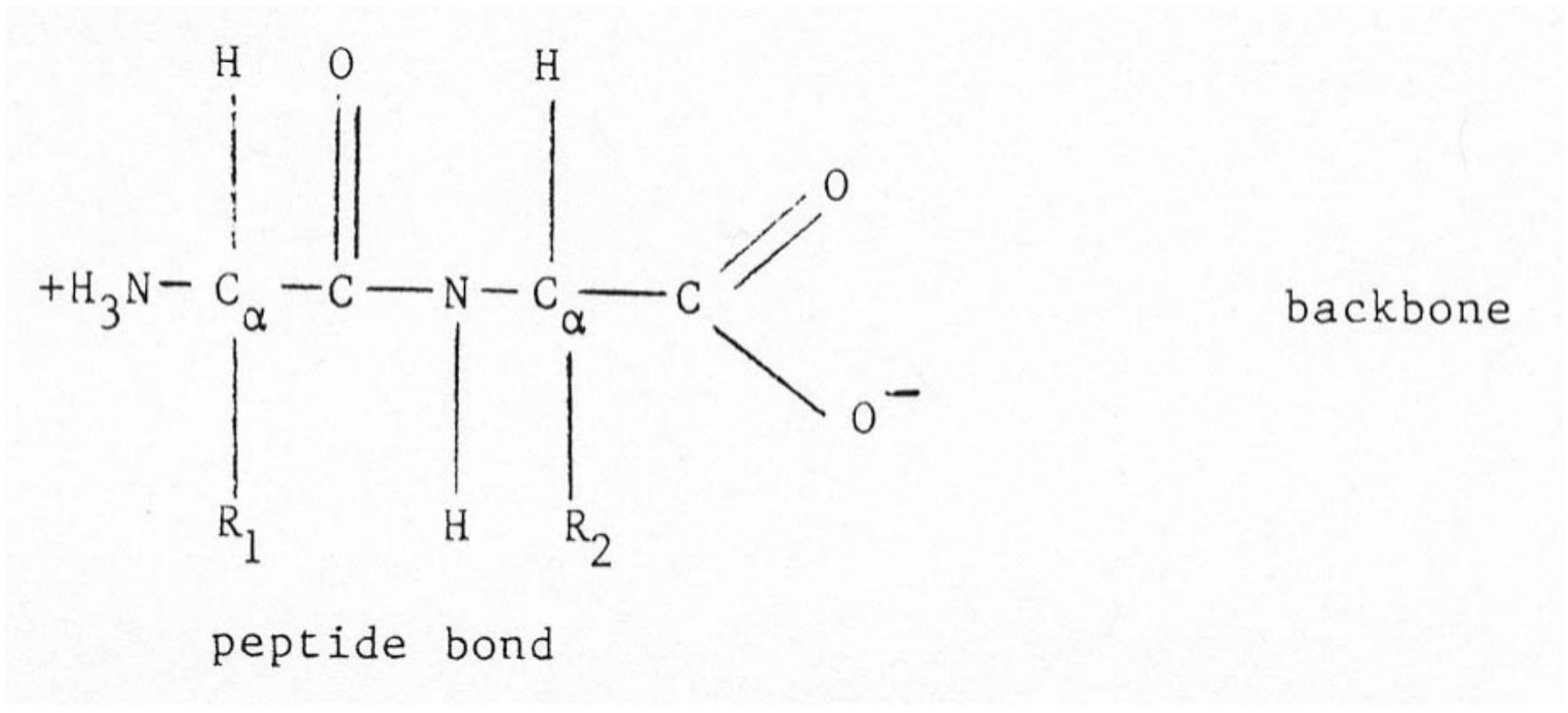


Fig. 4.3. A polypeptide chain formed from individual amino acids through peptide bonds.

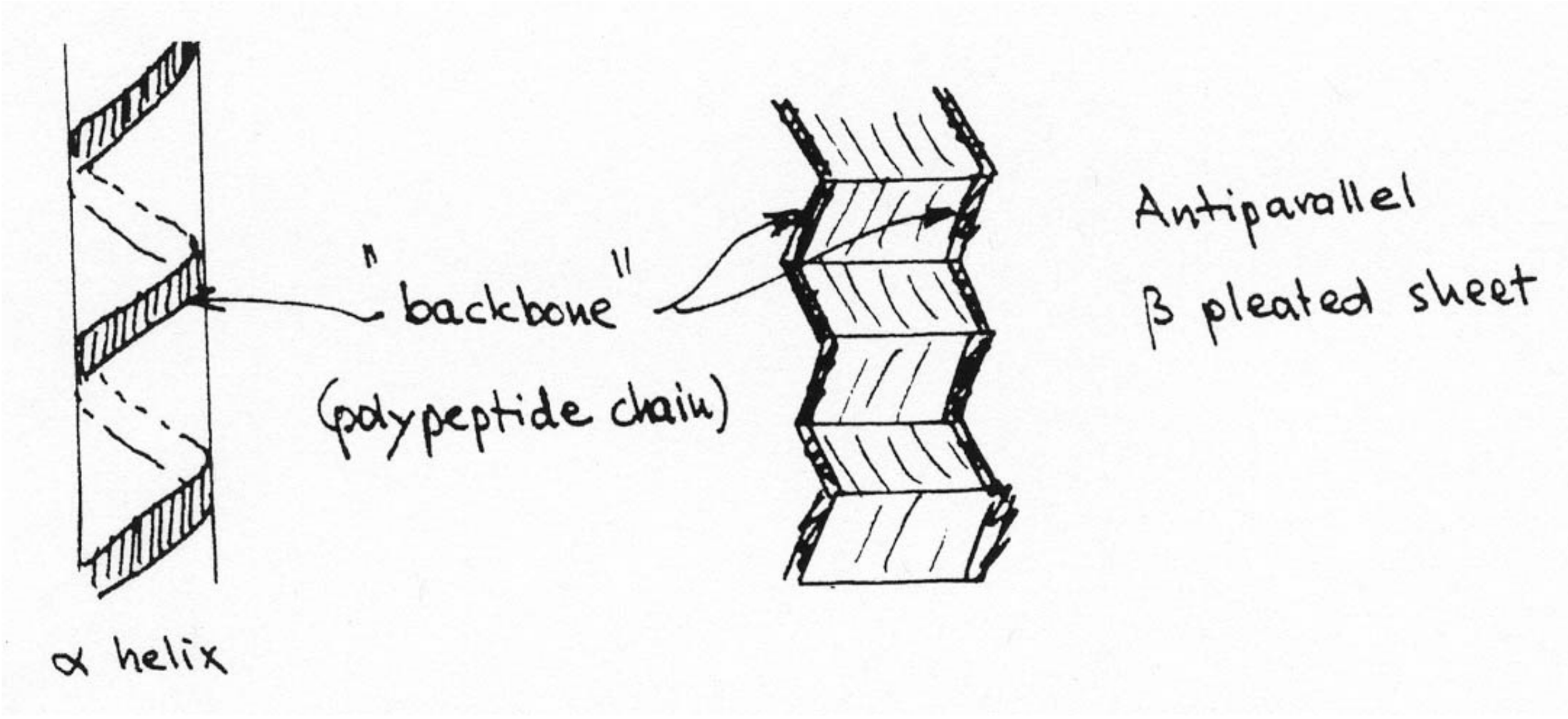


Fig. 4.7. Alpha helix and antiparallel beta pleated sheet.

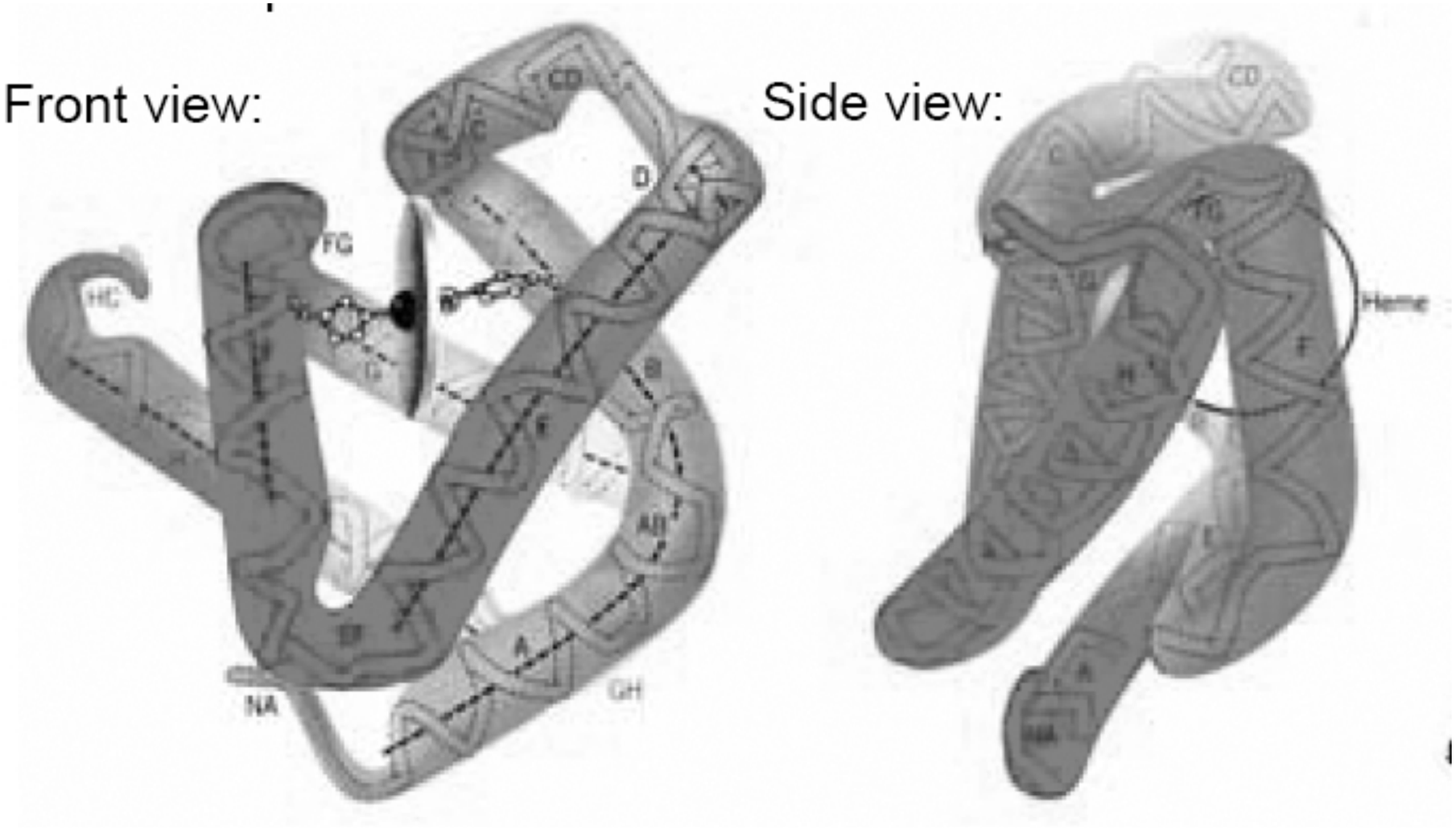
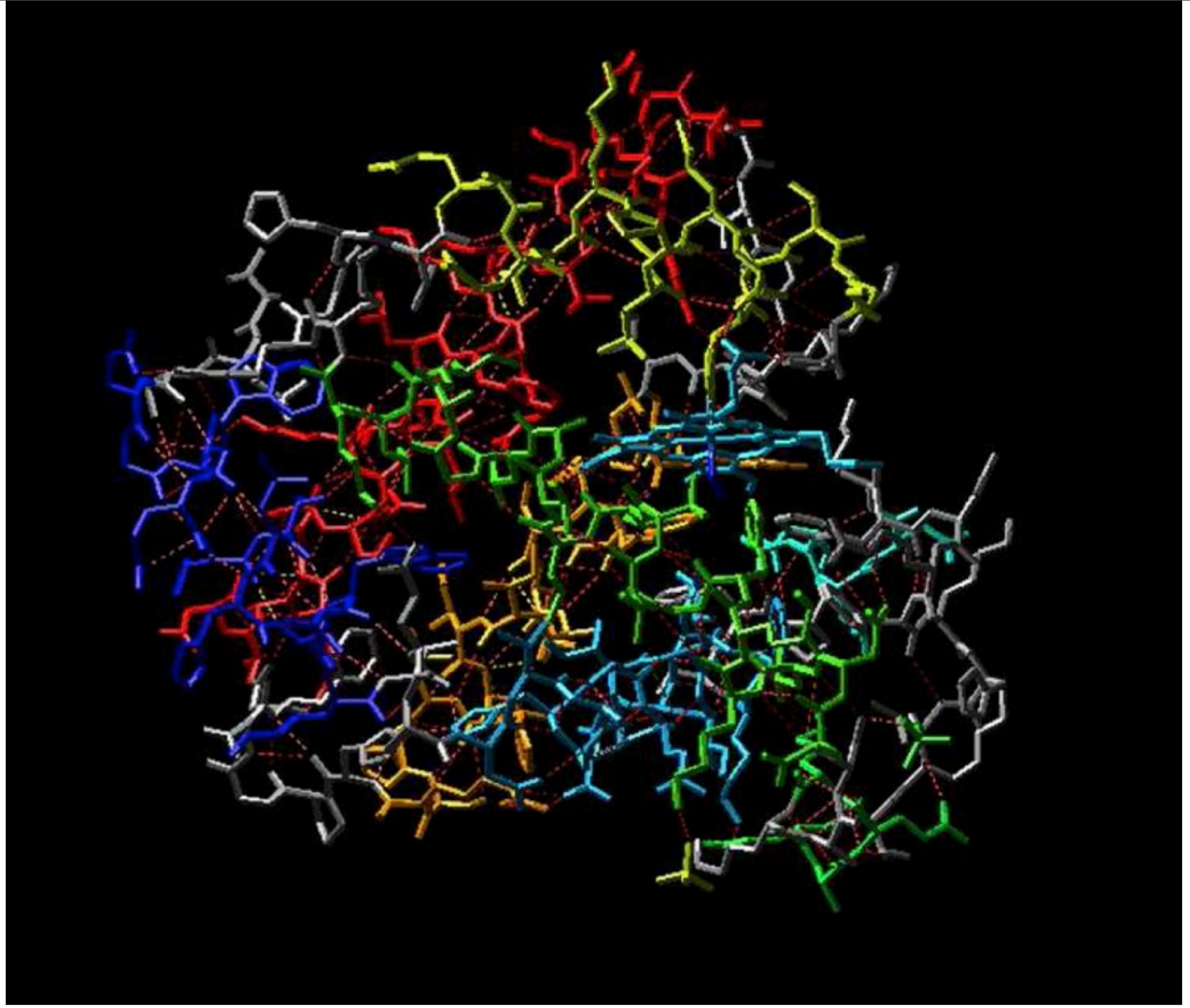


Fig. 4.11. The myoglobin molecule is built up from eight stretches of α helices that form a box for the heme group. Histidines interact with the heme to the left and right, and the oxygen molecule sits at point W. Helices E and F build the walls of the box for the heme; B, G, and H are the floor; and the CD corner closes the open end. After Dickerson and Geis [1].



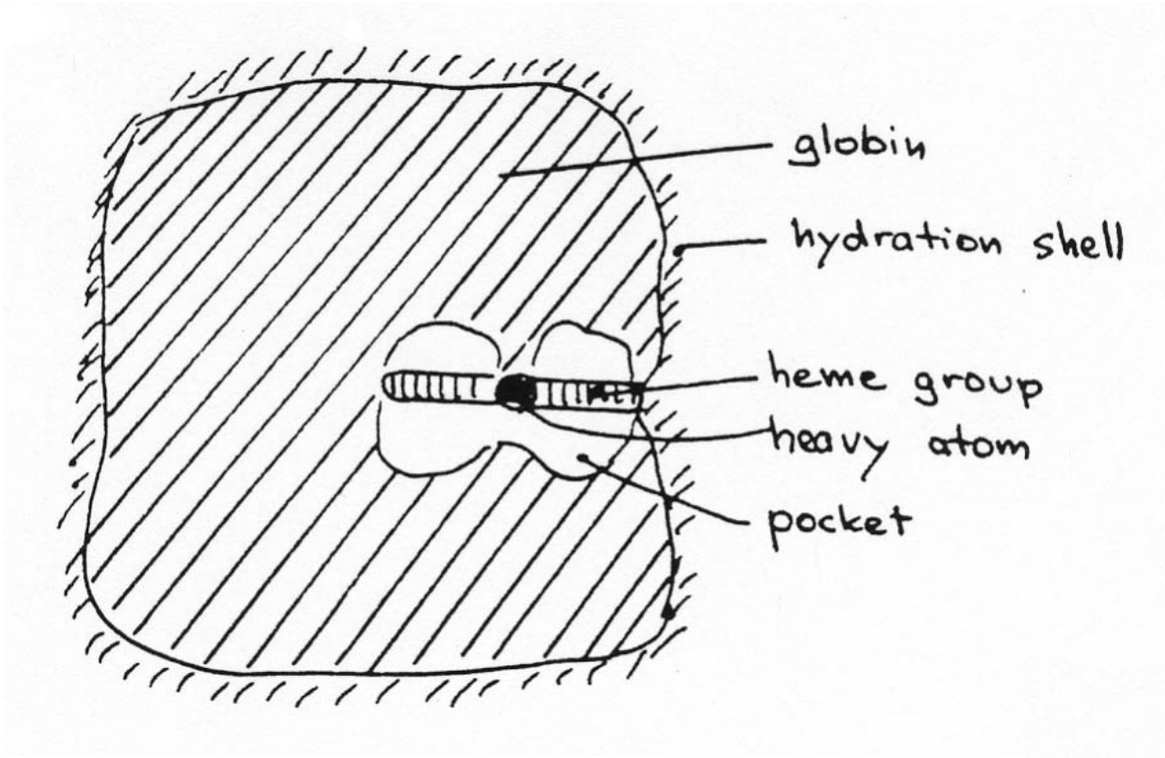


Fig. 4.8. A cross section through a heme protein.

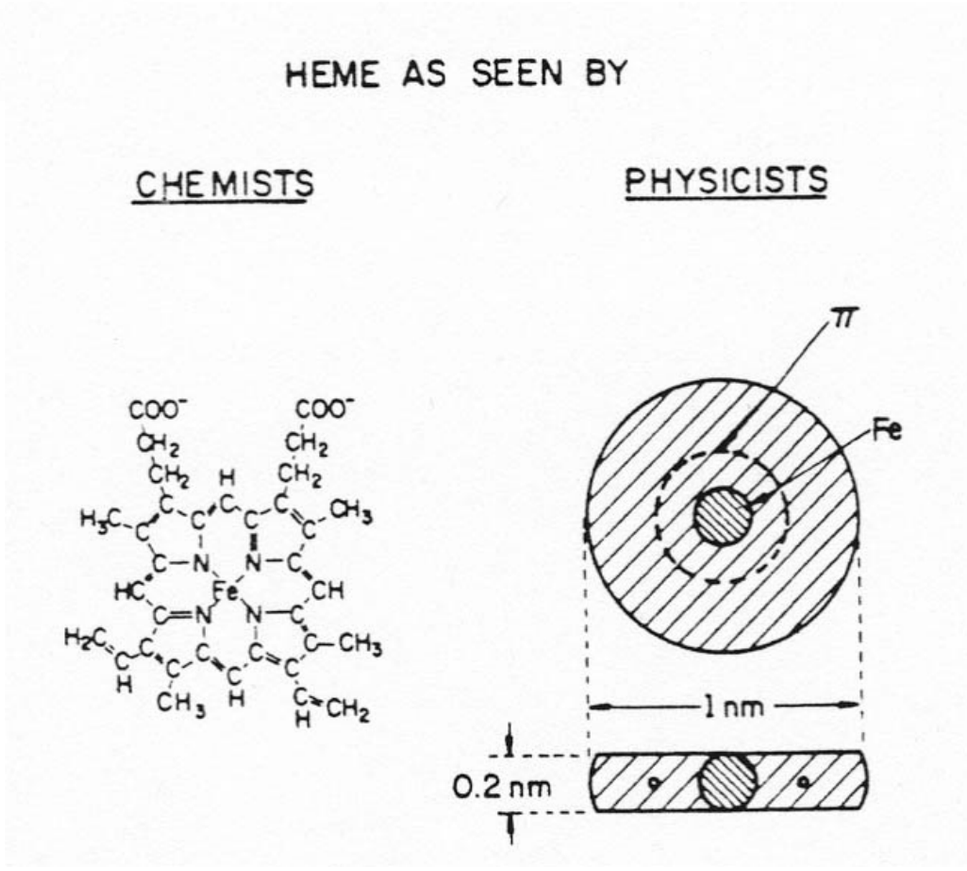


Fig. 4.9. The heme group as seen by chemists and by physicists. π indicates the pi-electron ring.

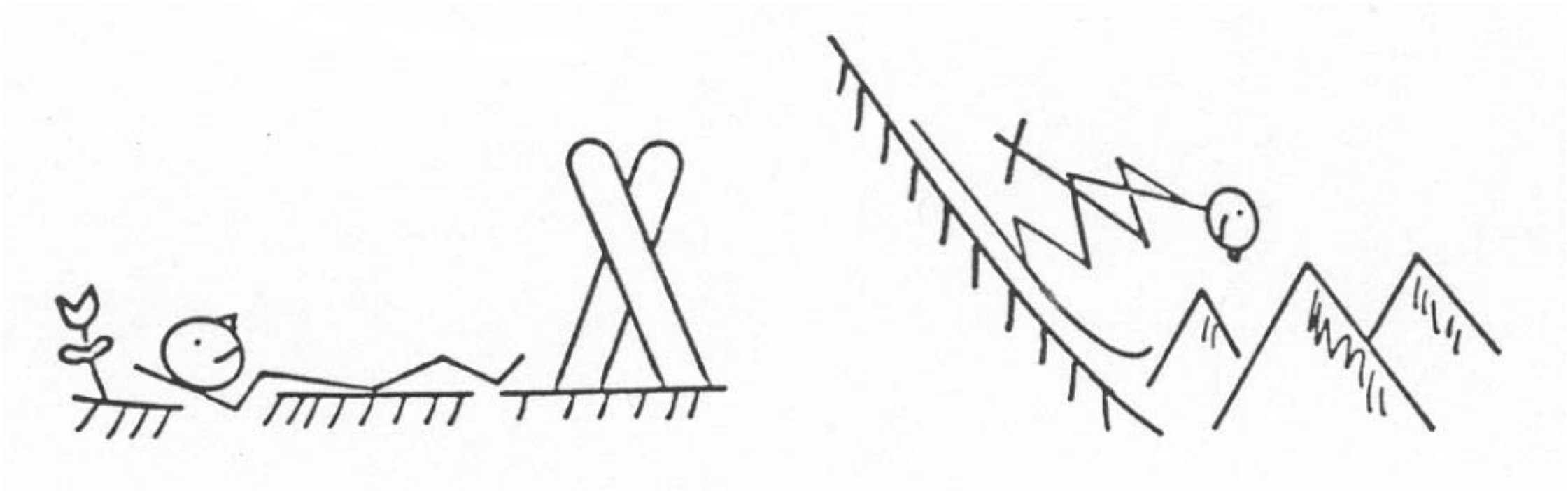
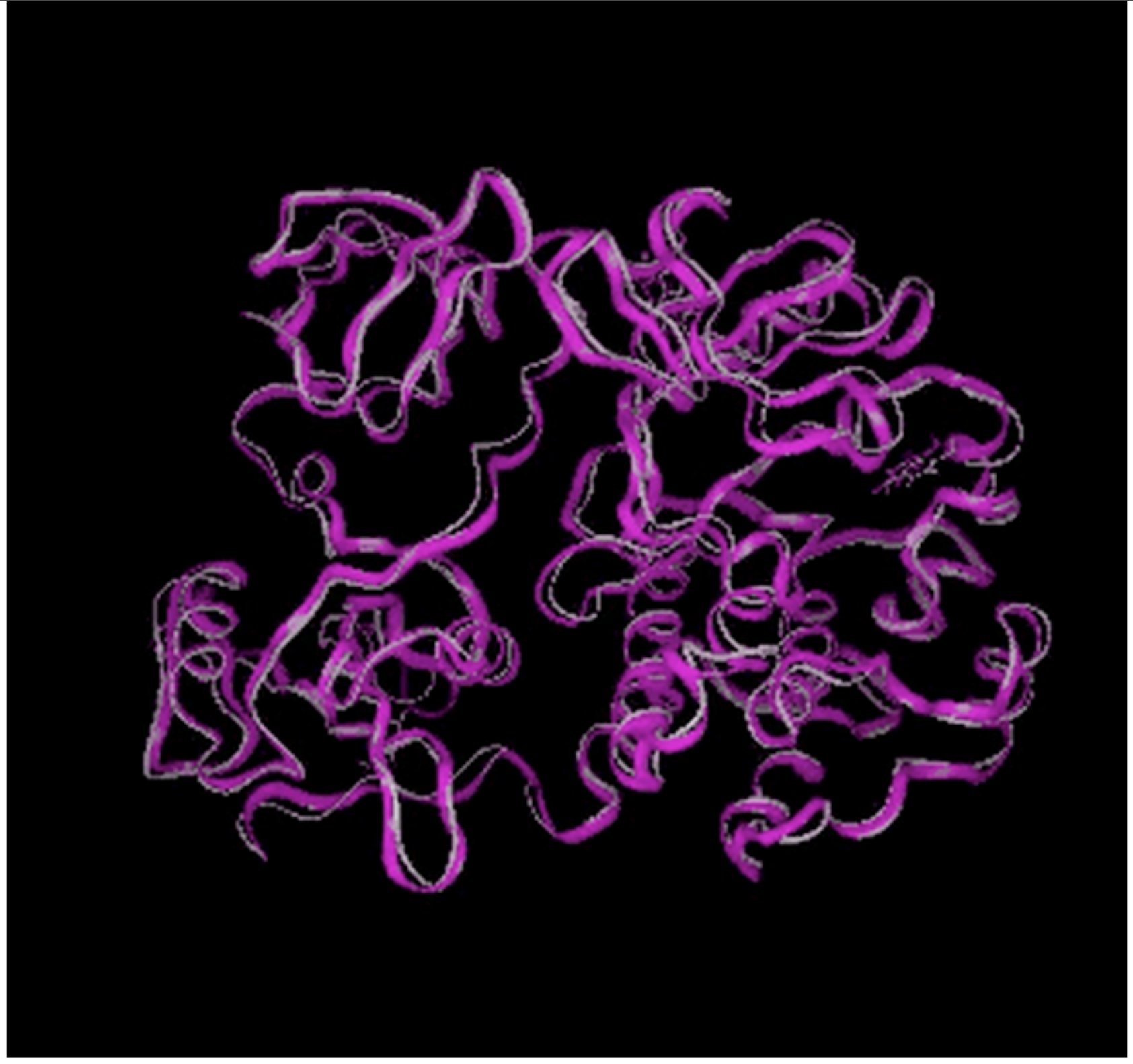


Fig. 11.2. A system at rest and in action.



Thermal motions jiggle the amino acids around about 0.1 nm at 300K

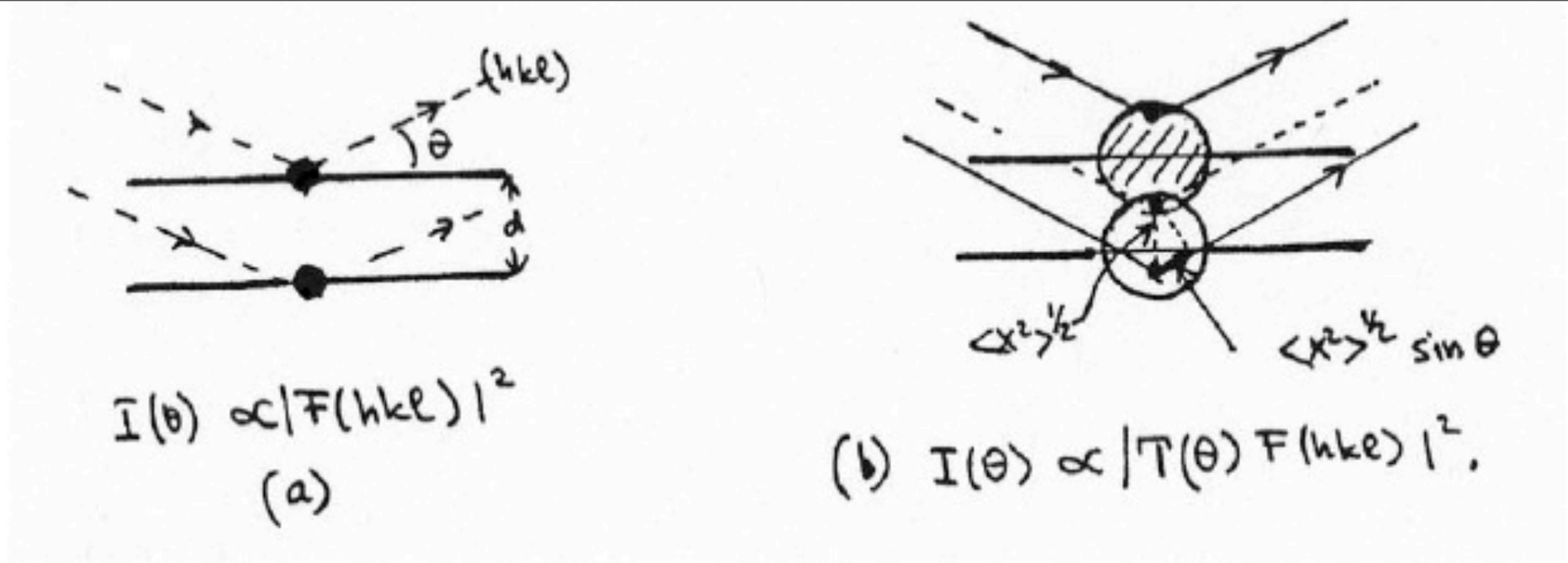


Fig. 11.8. (**a**) Diffraction from fixed atoms. (**b**) Atoms vibrate about their equilibrium position

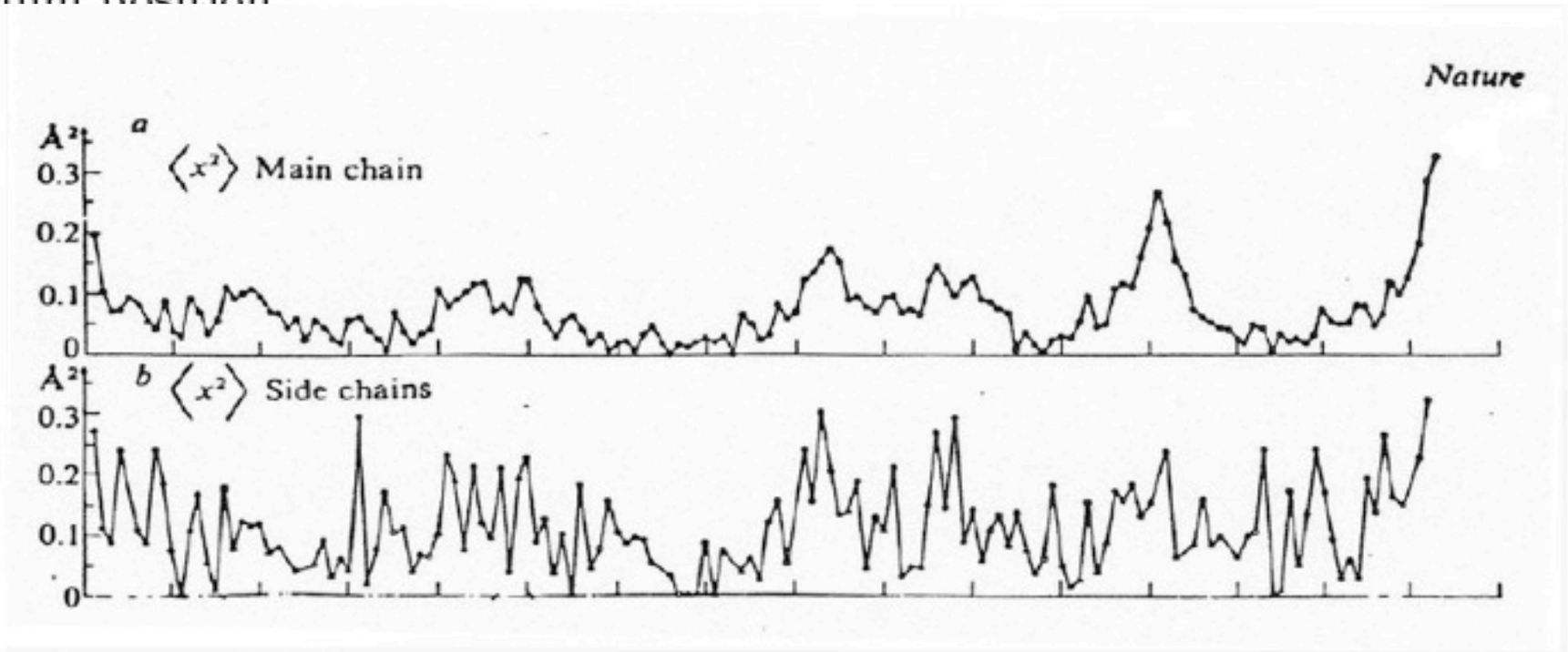
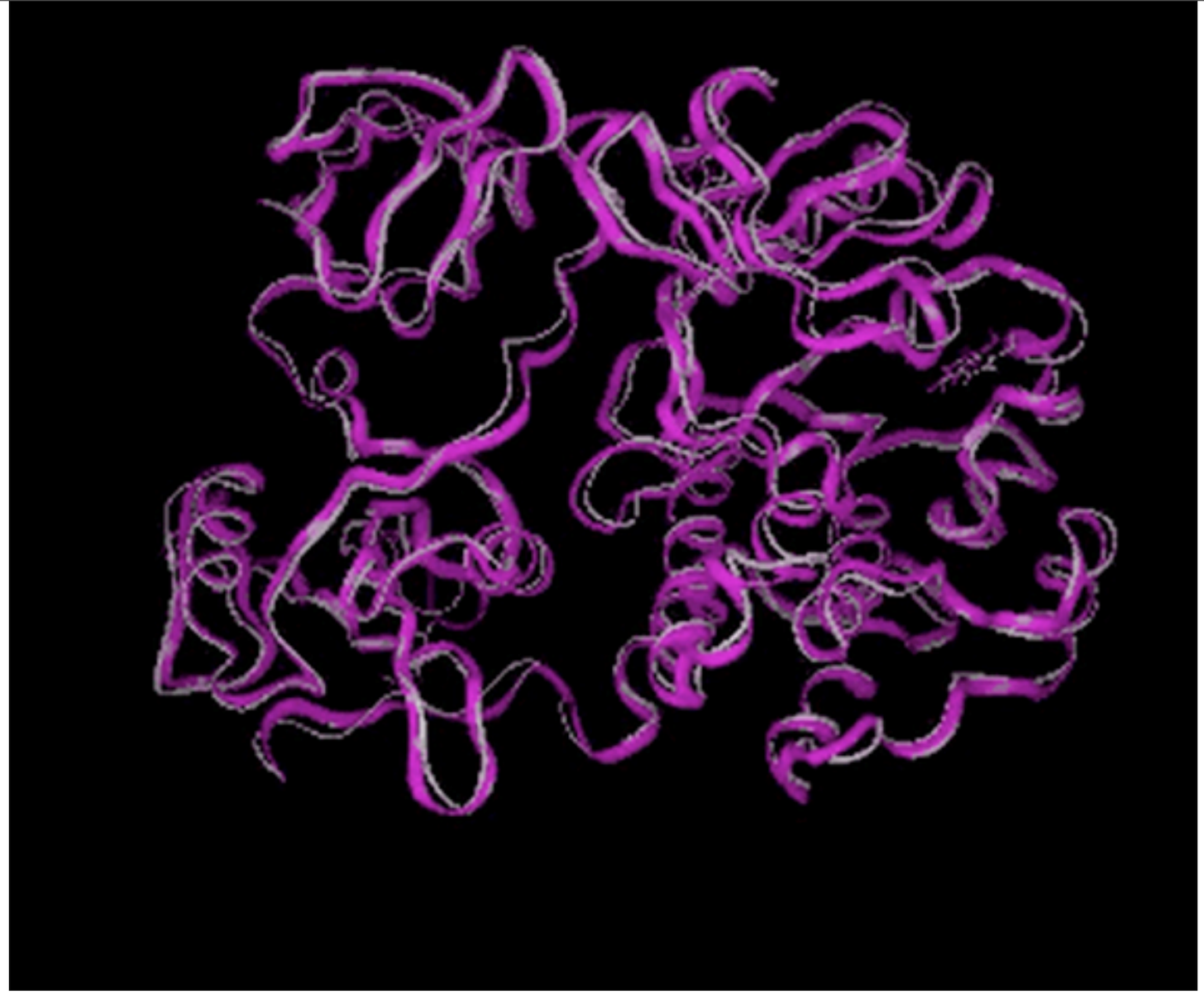
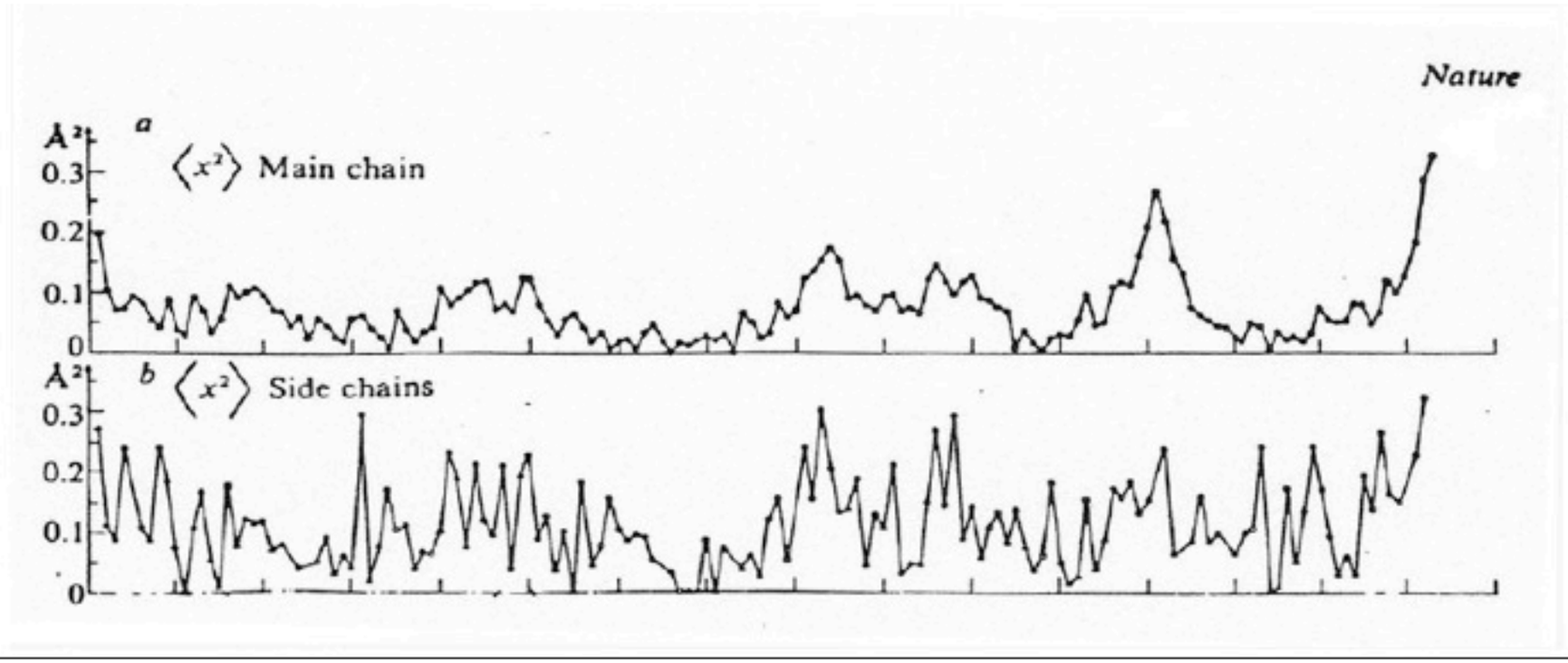
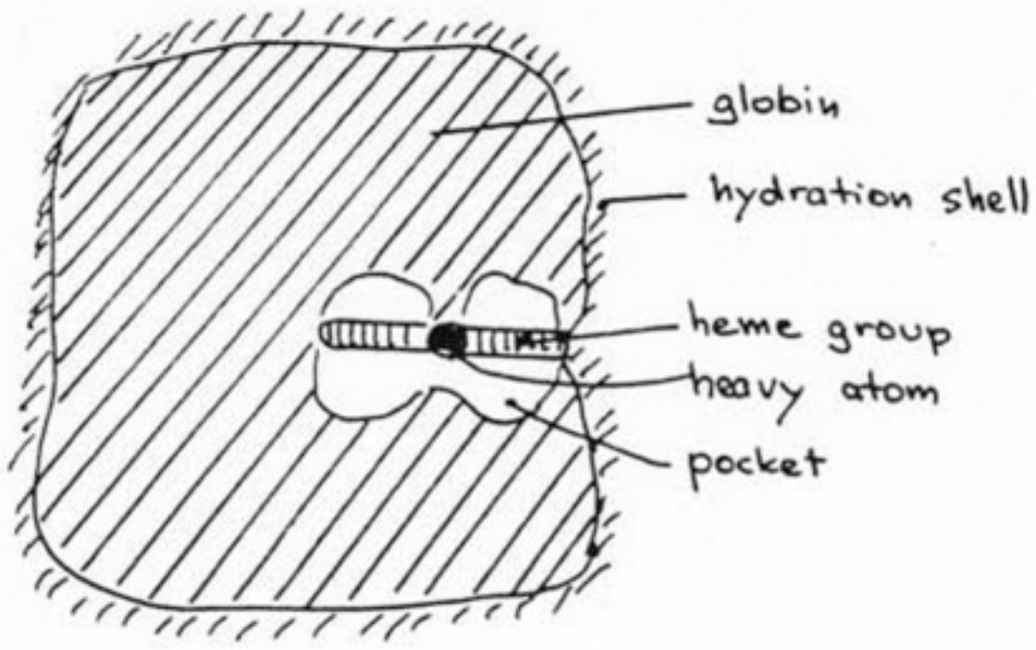
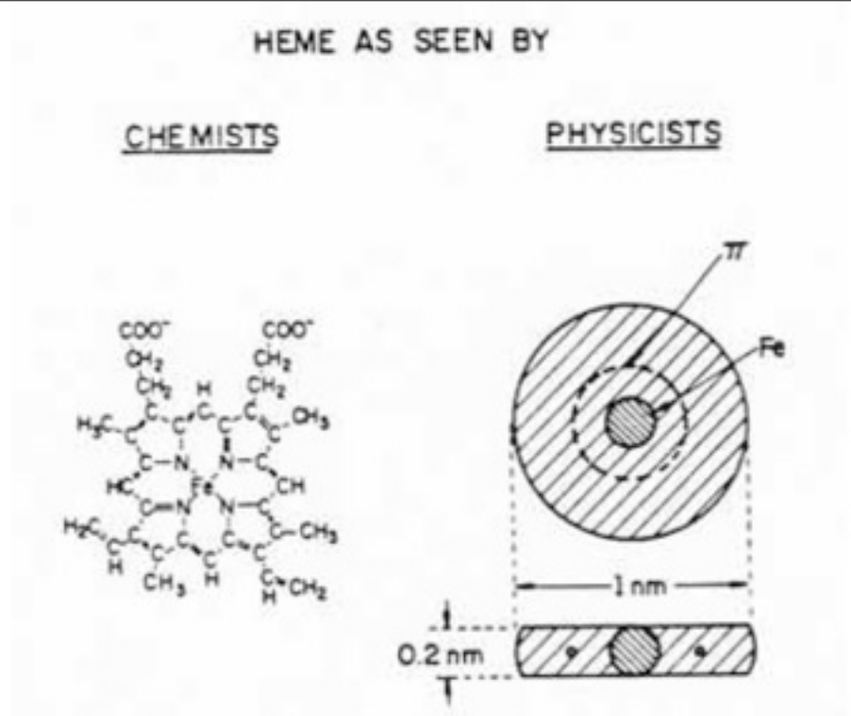


Fig. 11.10. The conformational and vibrational displacement $\langle x^2 \rangle_{cv}$ for myoglobin. The upper part gives the average values for the three mainchain (backbone) atoms C_α , C, N. The lower presents the largest x^2_{cv} in each sidechain [7].




Nature
a
Å²
0.3
0.2
0.1
0
⟨x²⟩ Main chain
b
Å²
0.3
0.2
0.1
0
⟨x²⟩ Side chains





$$\frac{dN}{dt} = -R(T)N[CO] \sim -k(T)N \qquad N(t,T) = N_o \exp[-k(T)t]$$

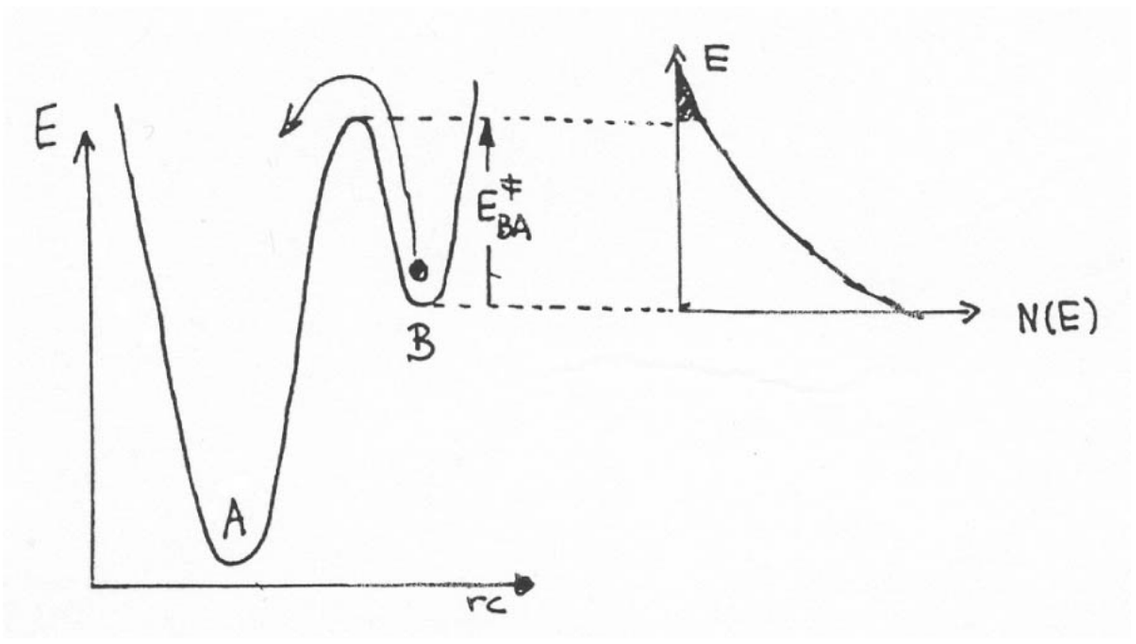
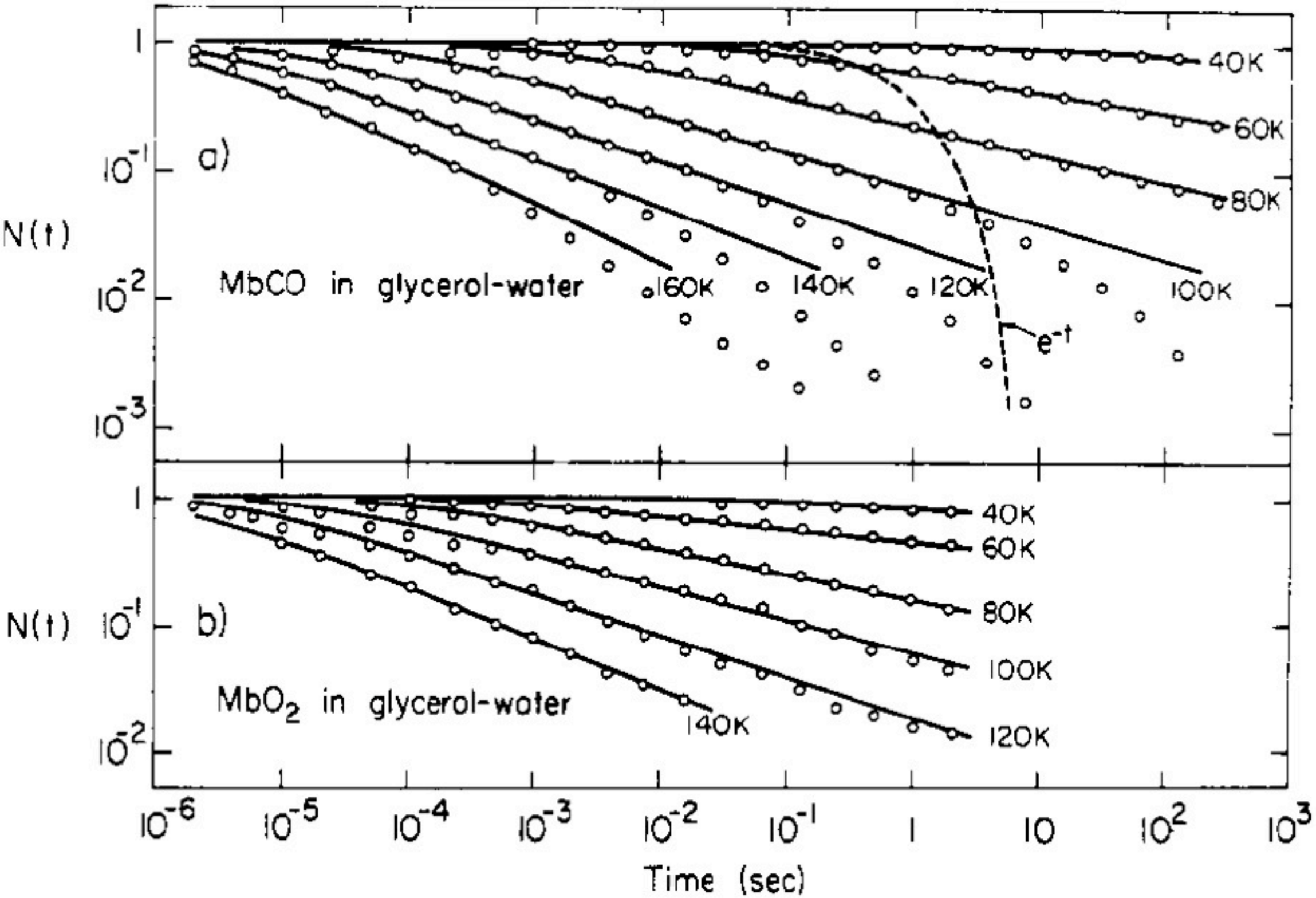


Fig. 13.6. Thermal transitions over the barrier between B and A. The Boltzmann distribution $N(E)$ vs E for the particles in well B are shown at right.

$$k(T) \sim [CO]R\exp[-H_{ba}/k_BT]$$

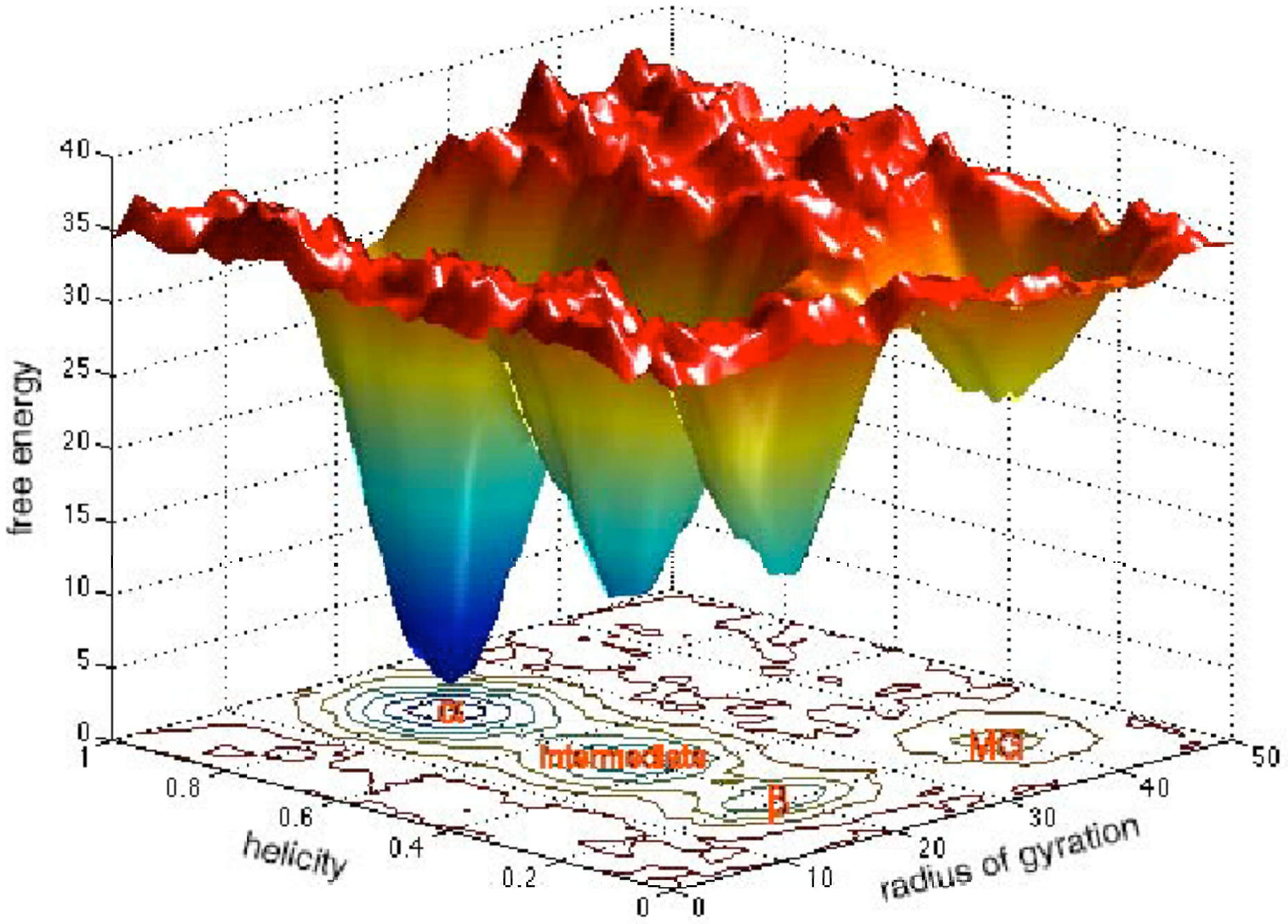


Power-law kinetics = distribution of conformations

Two things happen:

1) The system freezes into a distribution of states. There is no one protein conformation, but a free energy landscape of them.

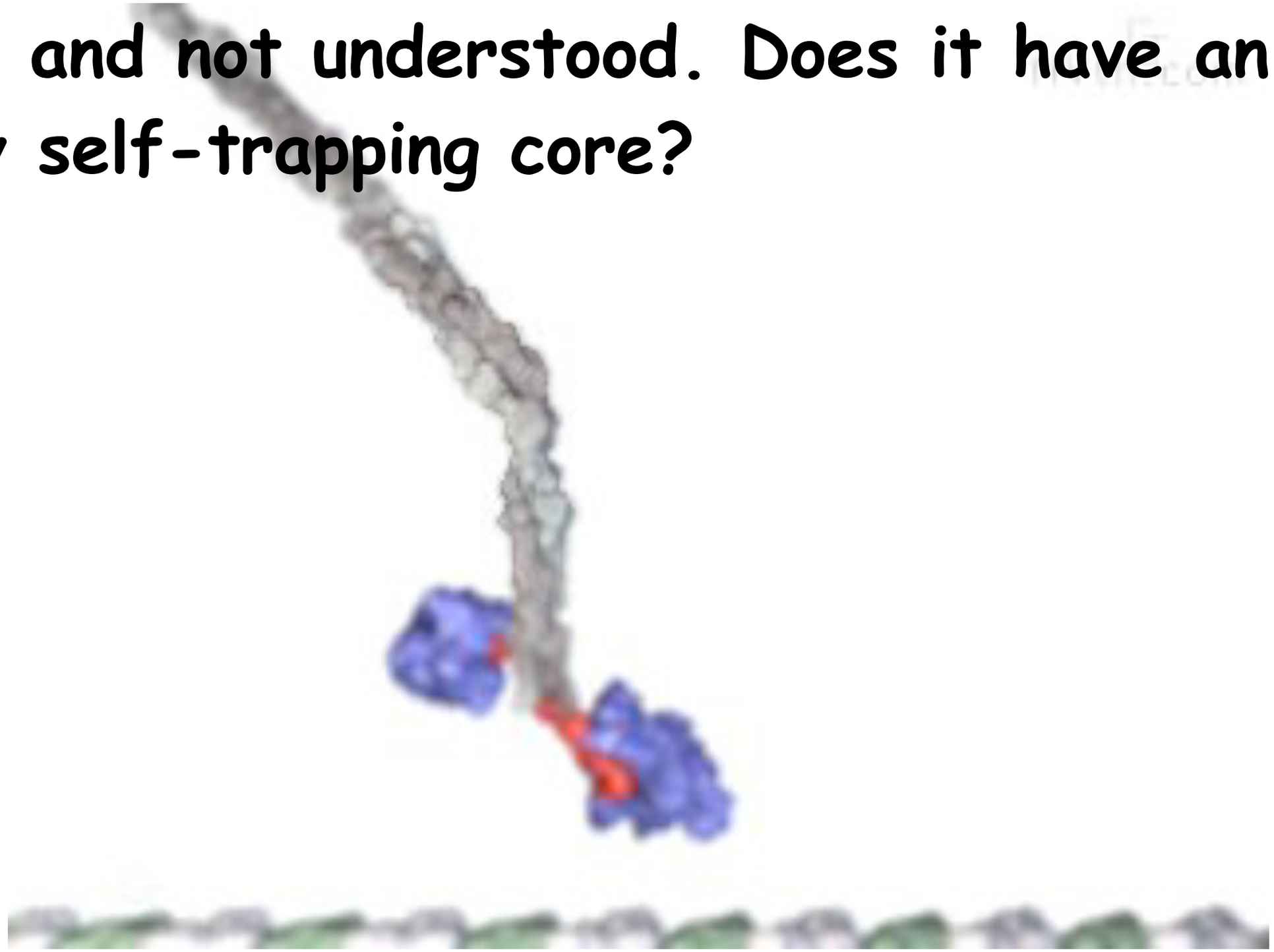
2) In the frozen distribution of states the "response function" of the system is a power law, NOT an exponential.


free energy
40
35
30
25
20
15
10
5
0
1
0.8
0.6
0.4
0.2
0
helicity
0
10
20
30
40
50
radius of gyration
α
Intermediate
β
MG

This work, done with Dr. CHAN Sui-ling remains my best cited work.

BUT.....it misses the true mysteries of what proteins actually DO.

Biological motors (kinesin): This is anharmonic, magic, and not understood. Does it have an energy self-trapping core?



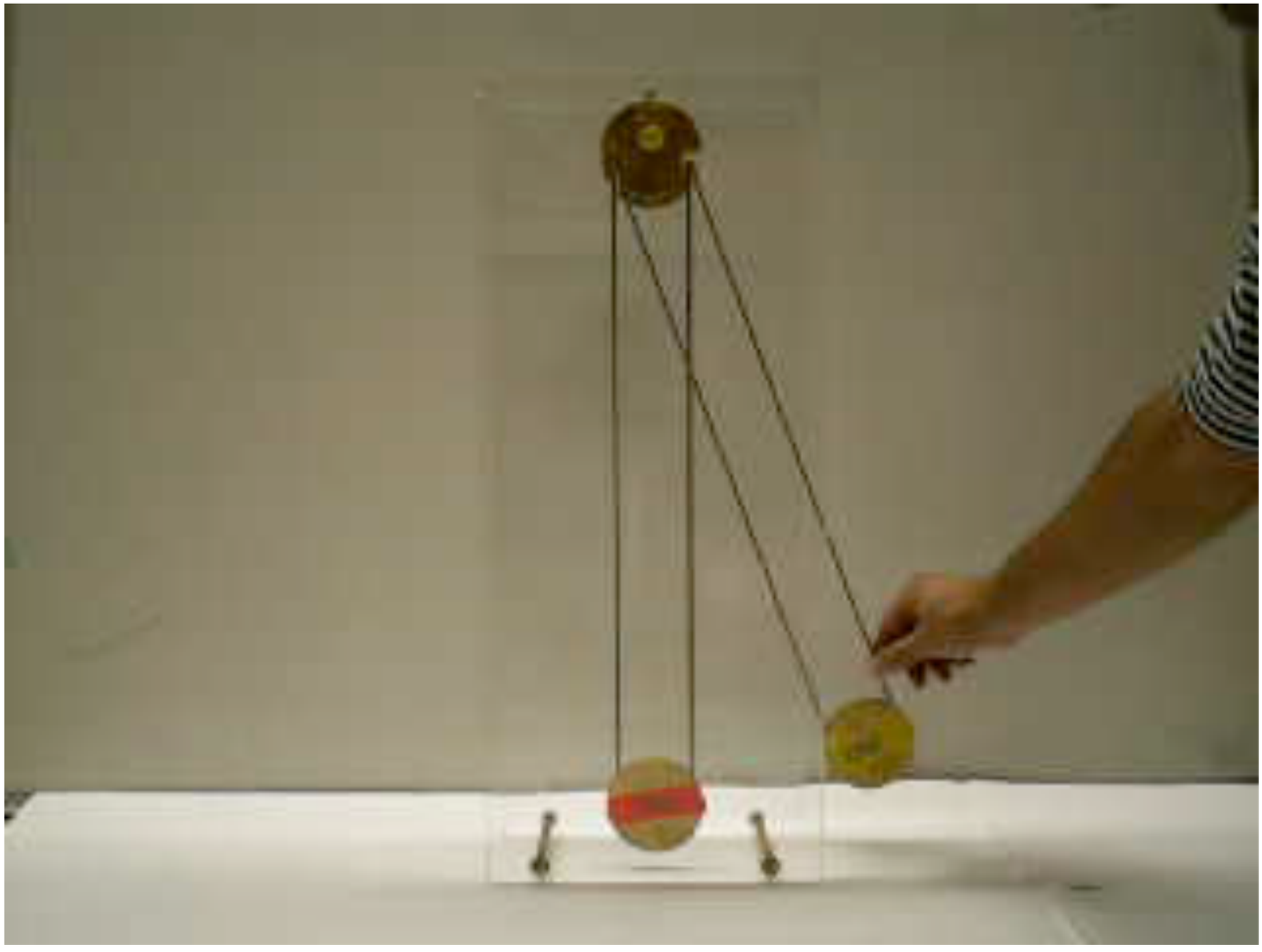
http://valelab.ucsf.edu/research/res_mec_overv.html

You all have seen this: weakly coupled pendula with normal mode beating:

$$H(\theta_1, \theta_2) = -mgL\cos(\theta_1) - mgL\cos(\theta_2) - \frac{1}{2}\kappa(\theta_1 - \theta_2)^2$$

You all have seen this: weakly coupled pendula with normal mode beating:

$$H(\theta_1, \theta_2) = -mgL\cos(\theta_1) - mgL\cos(\theta_2) - \frac{1}{2}\kappa(\theta_1 - \theta_2)^2$$



But have you seen this:

But have you seen this:

This is about self-trapped energy states in protein.

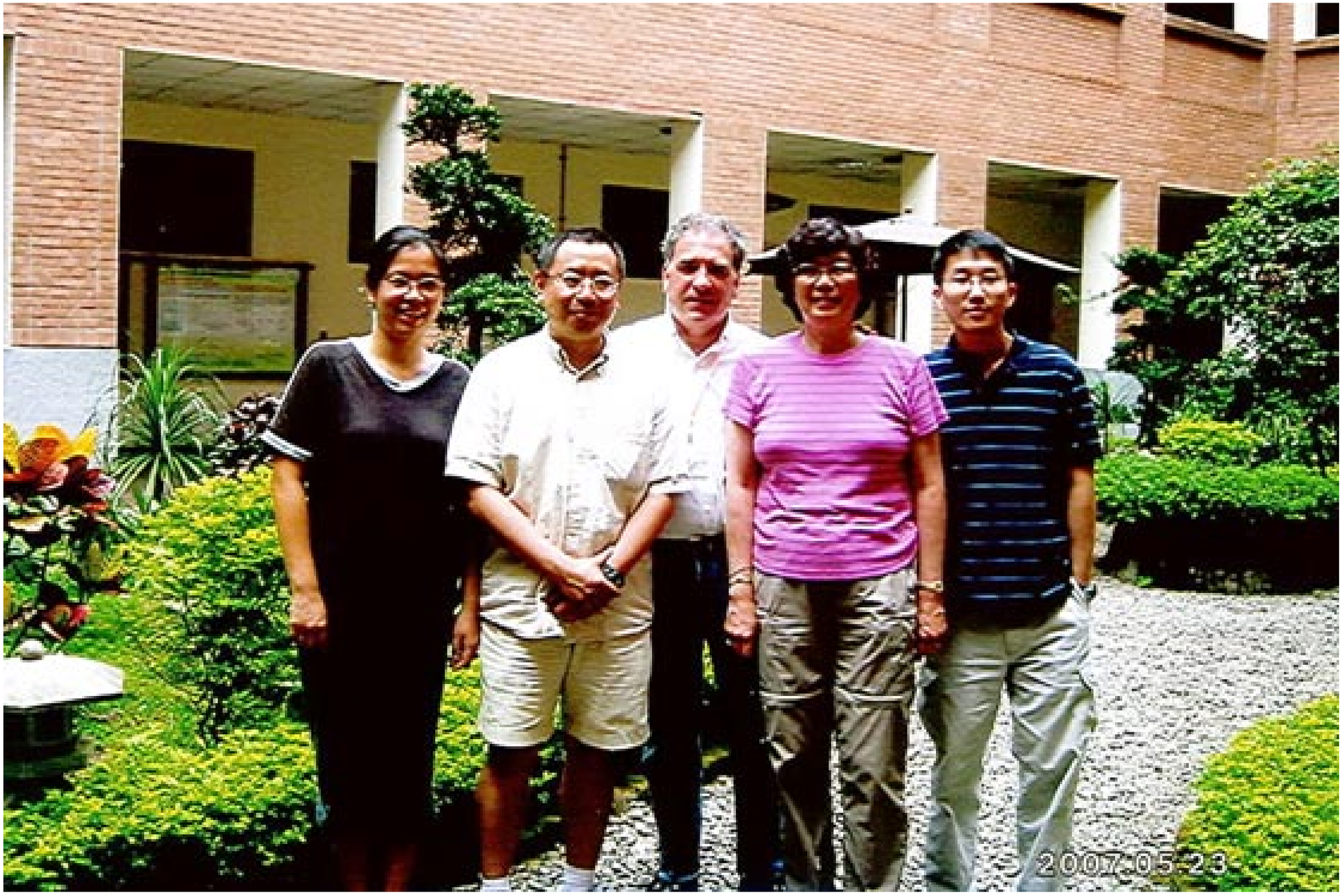
Self-trapping of energy is a non-linear process where energy localizes in a particular mode.

There is a long history coming from condensed matter theory about quantum self-trapping of energy in coupled anharmonic systems.

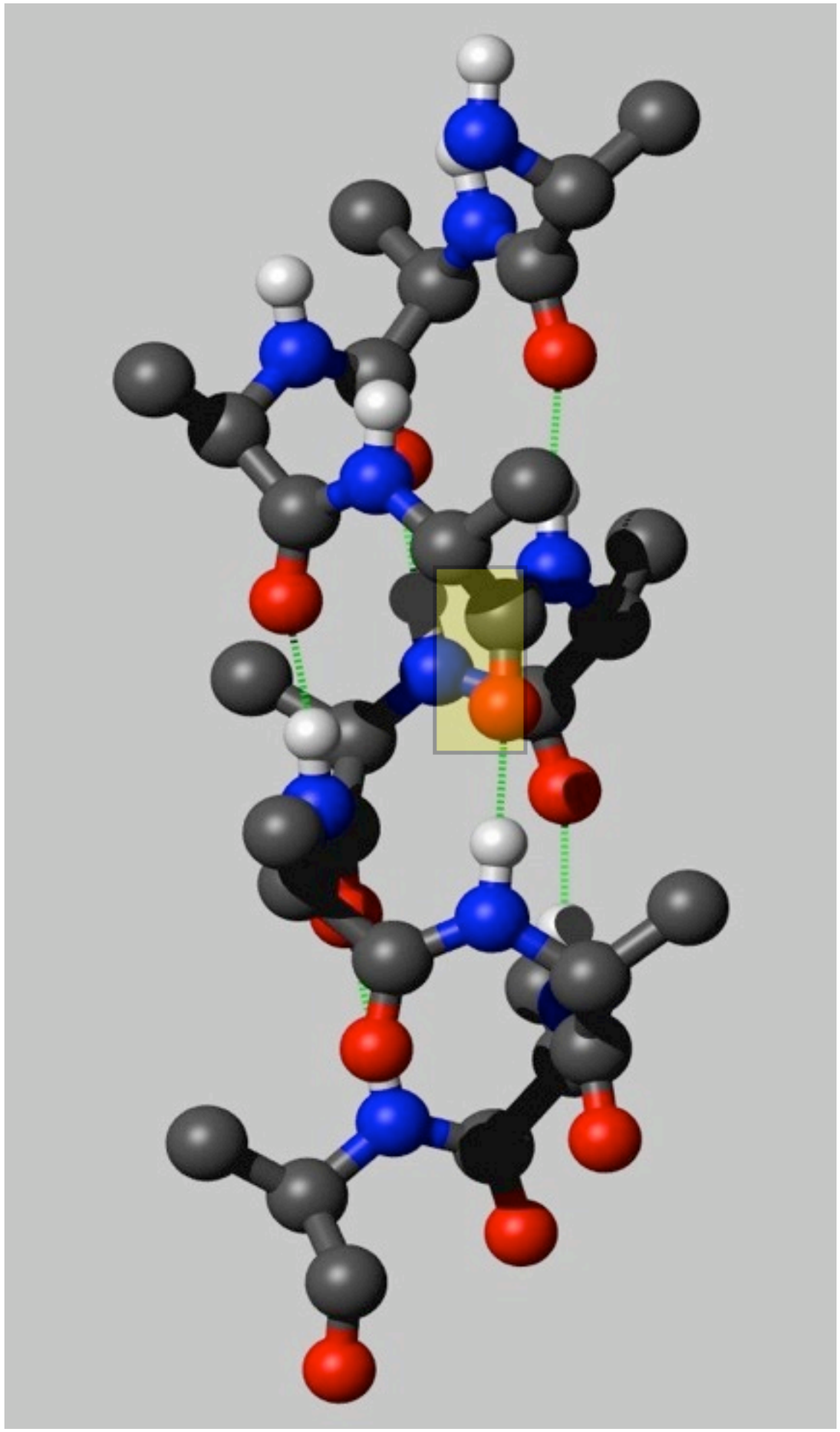
Davydov, A.C. Scott tried to take this into protein dynamics, and I got trapped..

Monday,13 July, 09

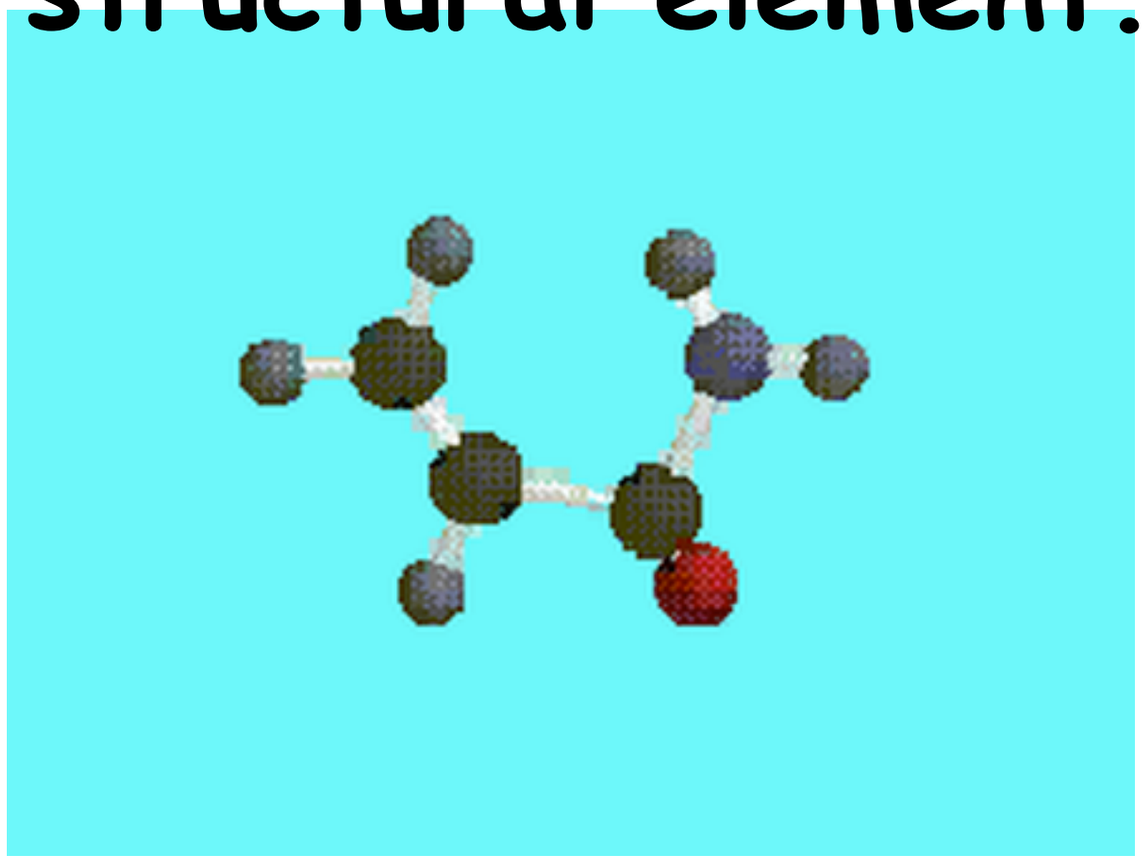
Surfing China's Qiantang River Tidal Bore, Hangzhou - September, 2008



A farewell to Arms....the Free Electron Laser wars to find solitons.

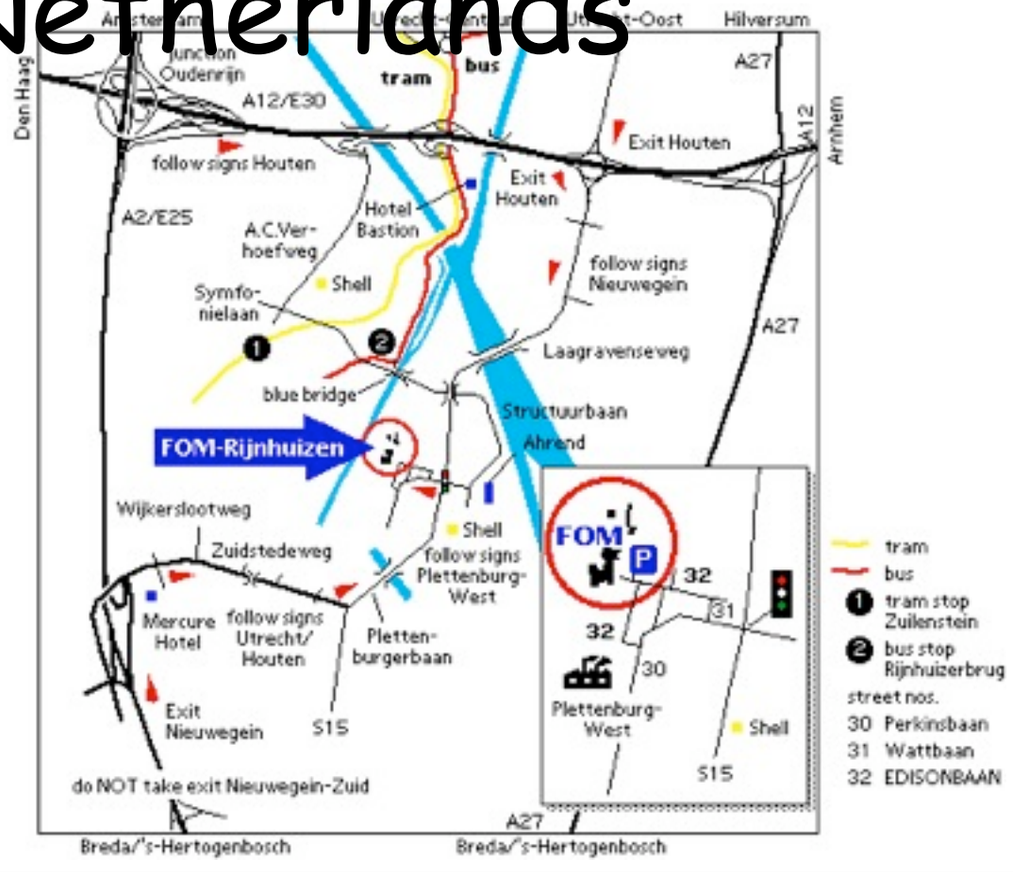


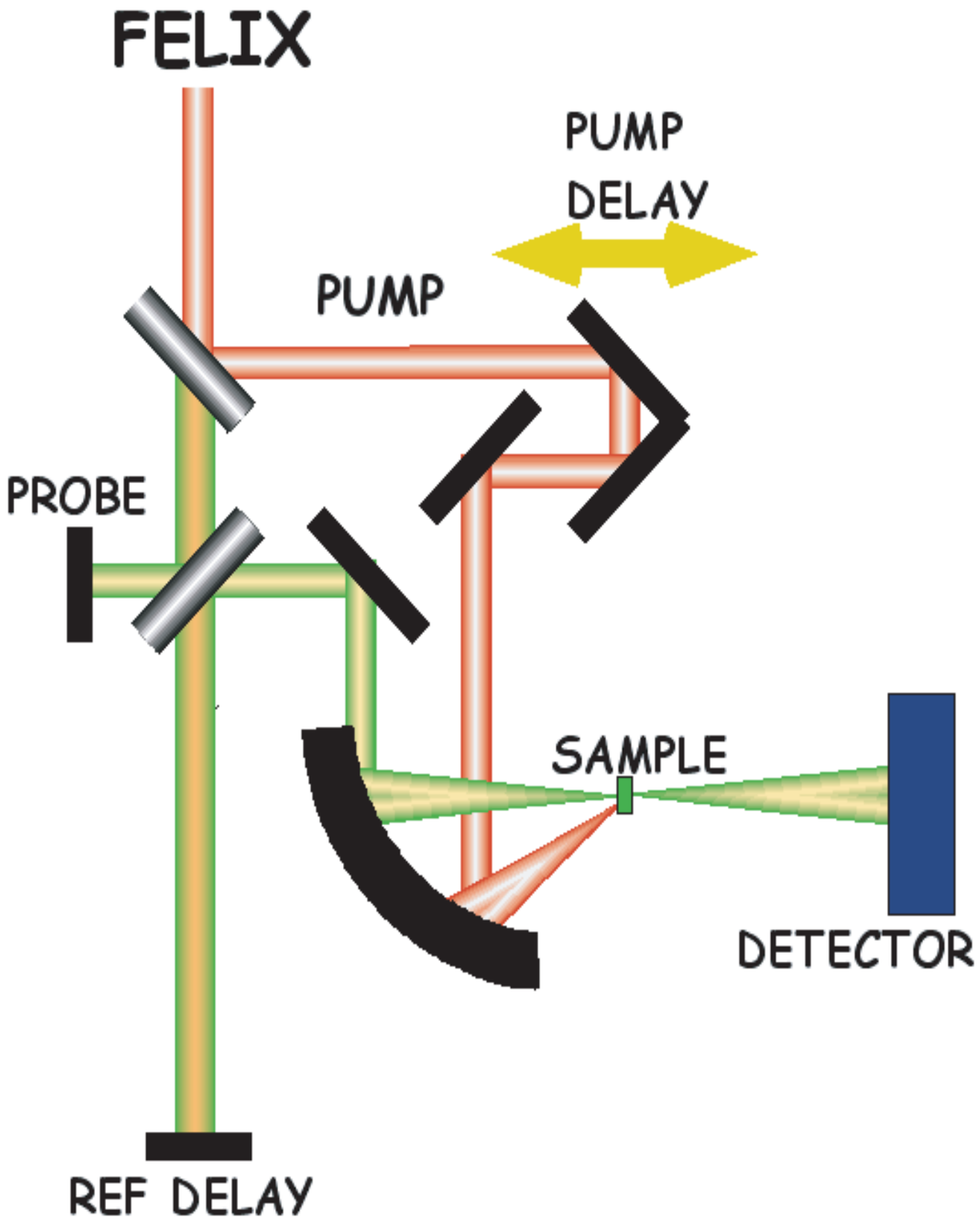
The alpha-helix is an important structural element.



The amide-I mode is a C=O stretch coupled to N-H, 6 um (1650 cm-1)

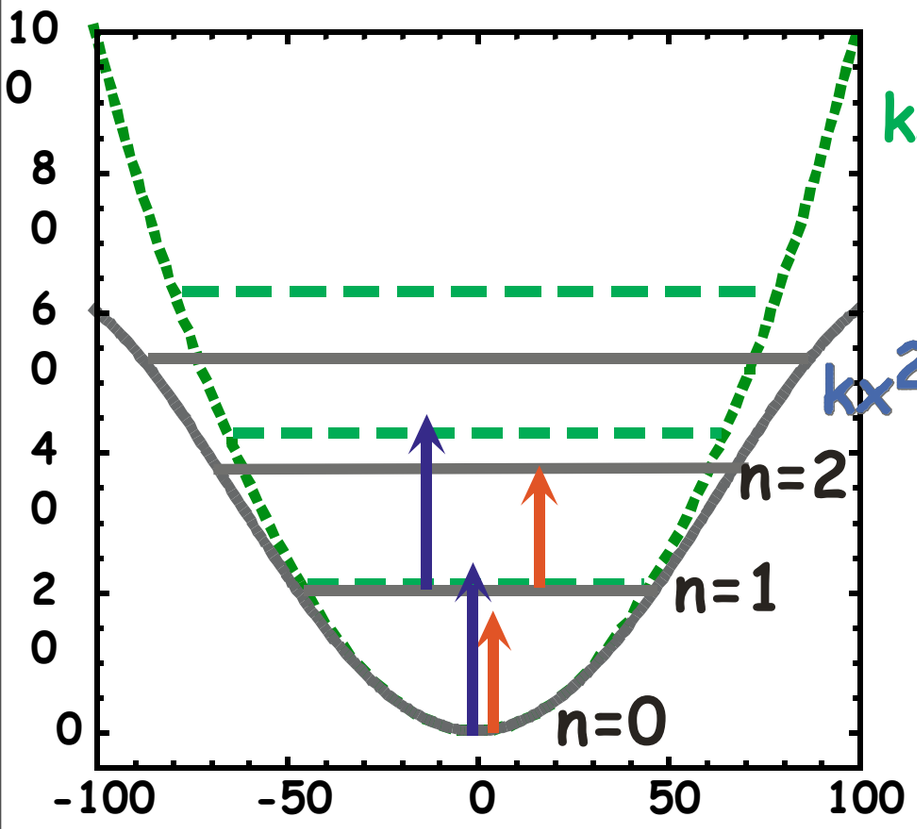
FELIX: the marvelous IR-FIR (3 microns to 100 microns!!!) Free Electron Laser near Utrecht, the Netherlands

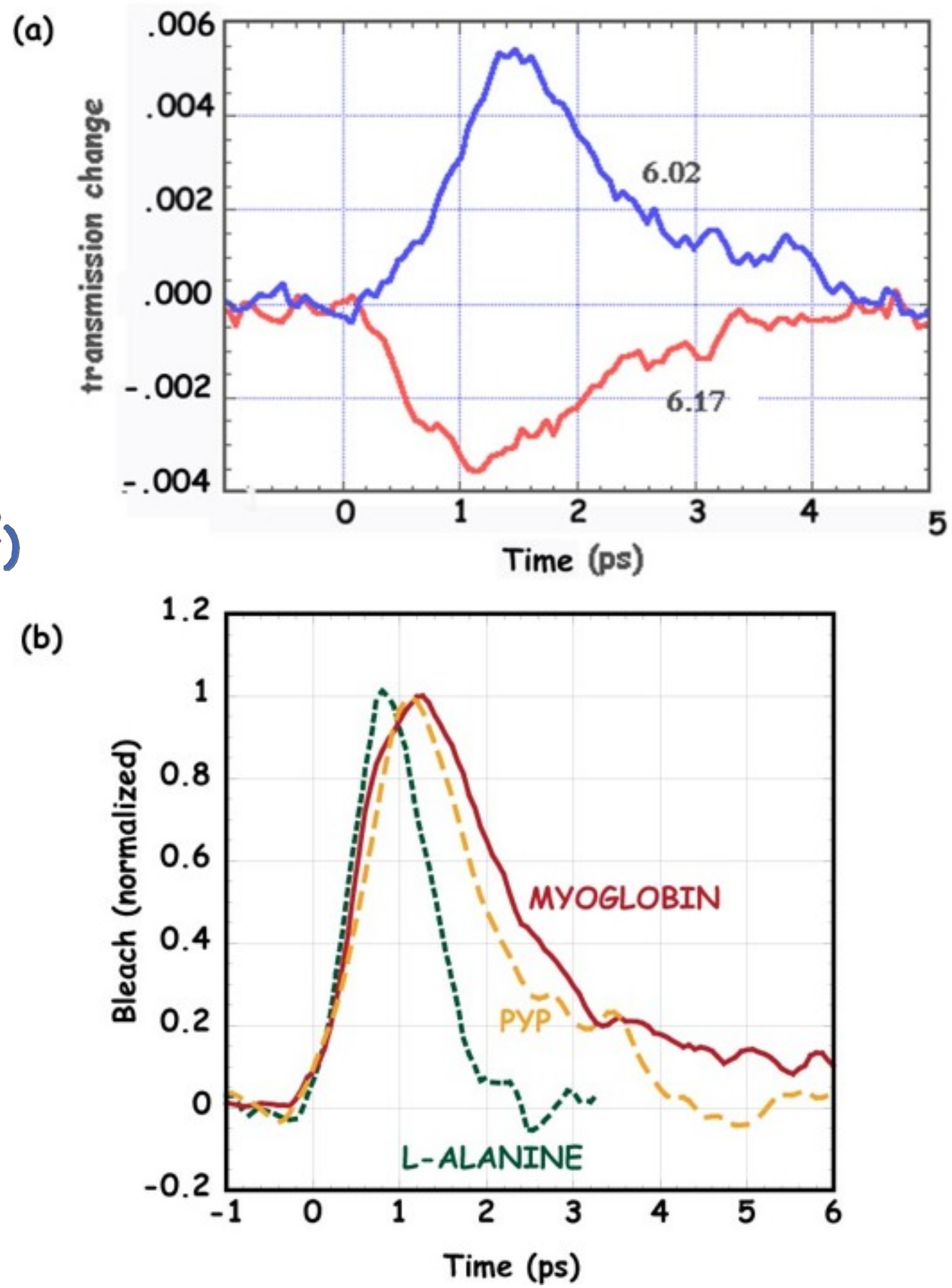


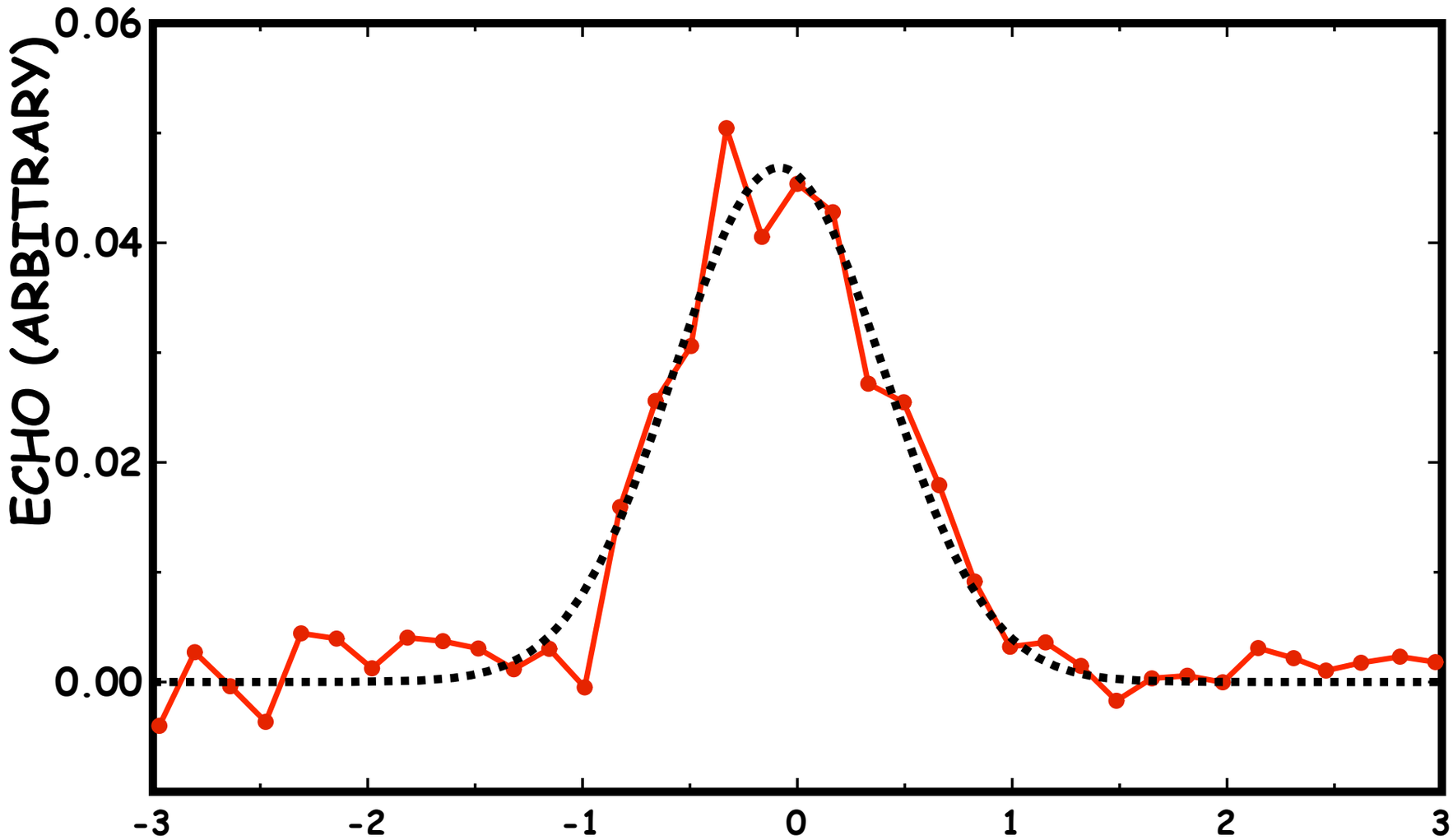


Pump-probe experiments are direct probes of anharmonic effects: a harmonic system cannot be saturated. In the IR, you can selectively pump modes.

Small, fast, boring.



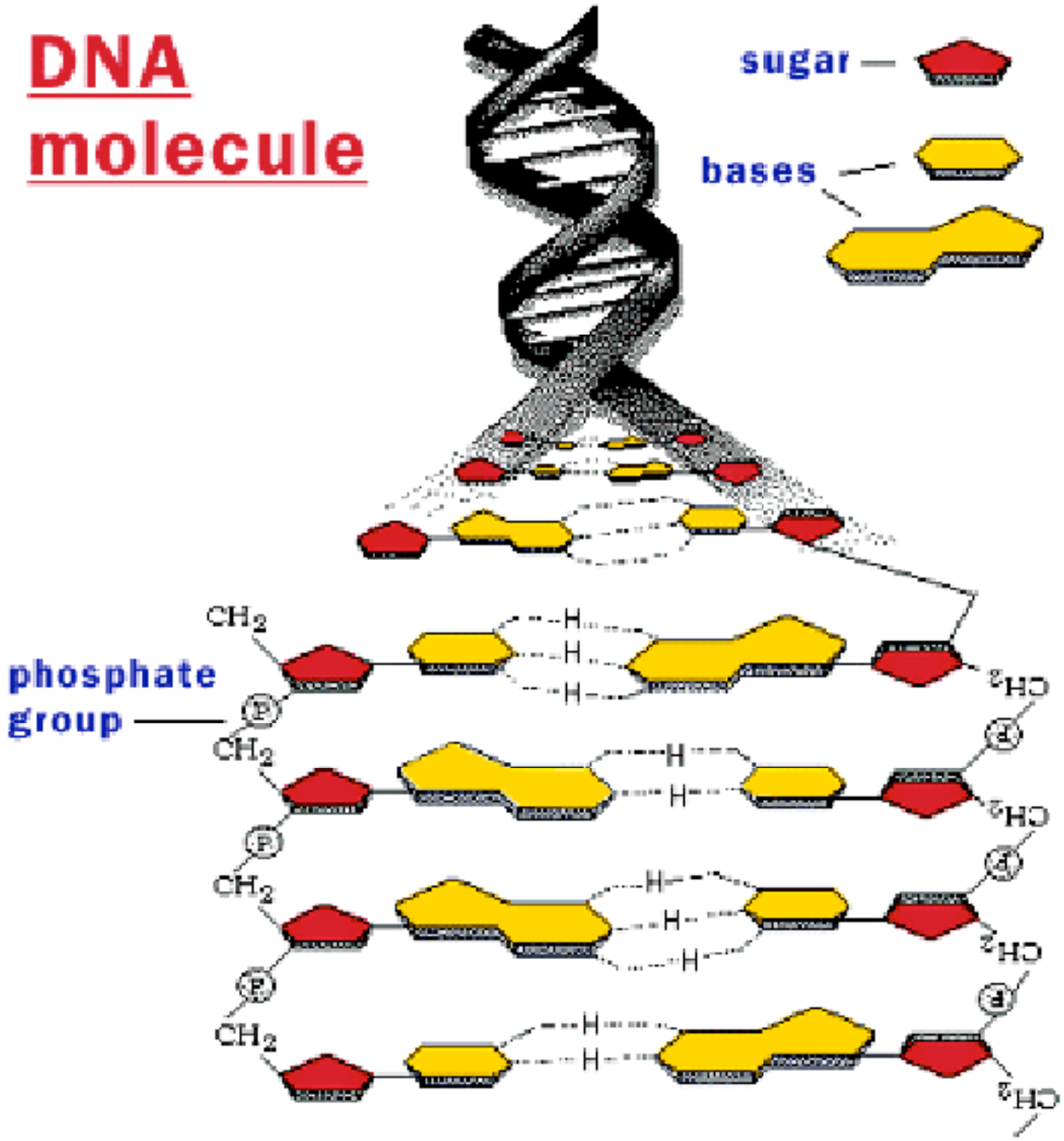




NO long-lived wave like the "silver dragon" in Hangzhou!

2) The softness of DNA and why it matters.

DNA molecule



DNA contains the code for you. There are about 3 billion basepairs in the human genome, or about 1 meter of DNA.

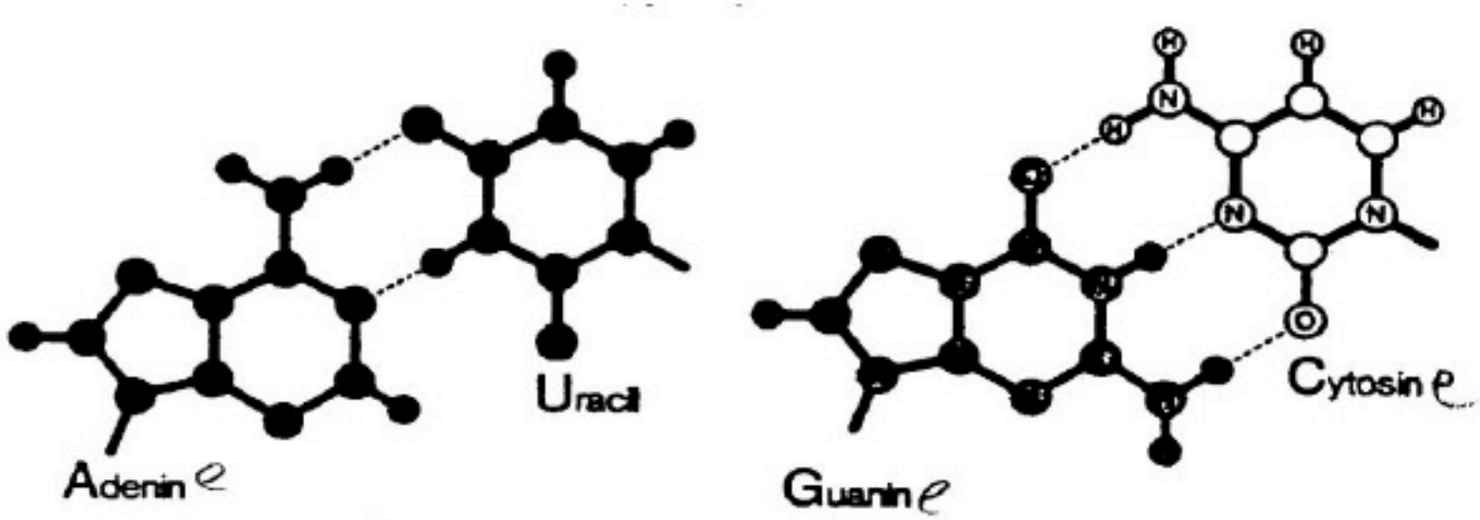
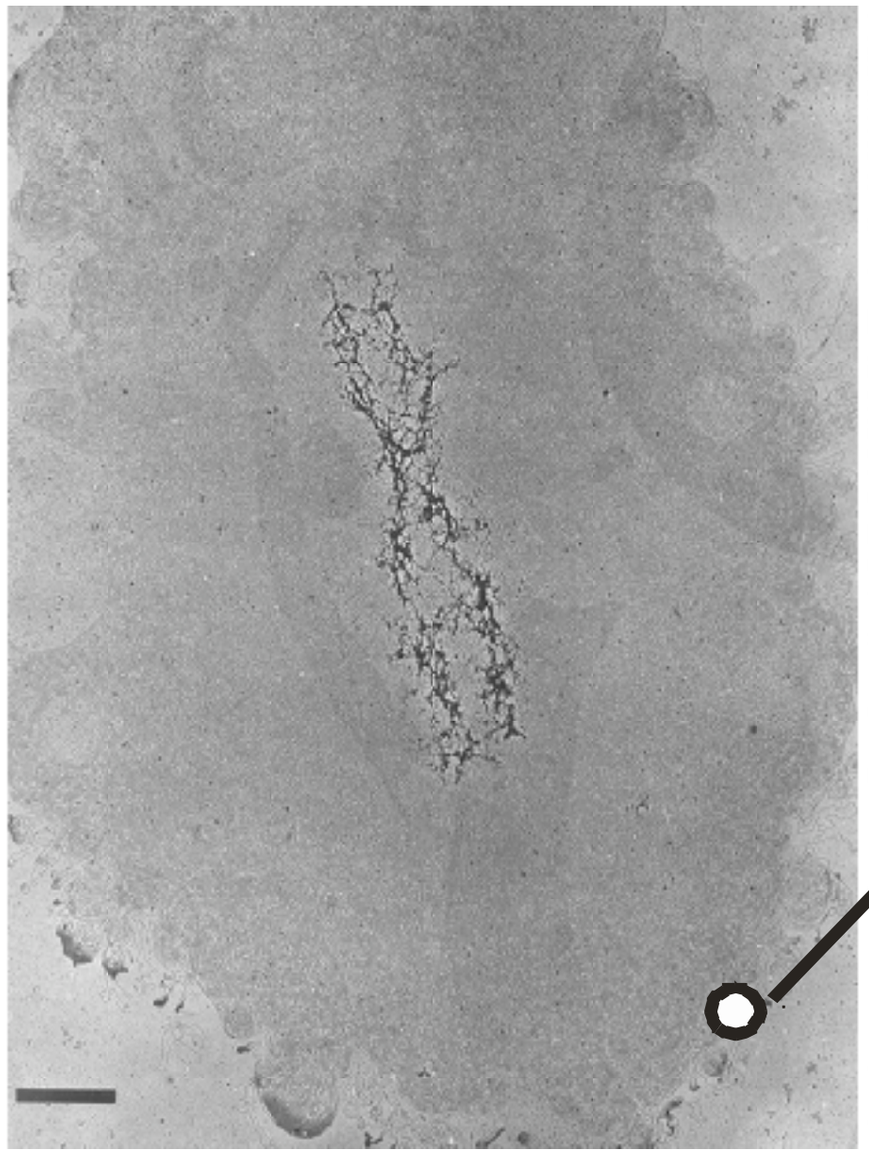
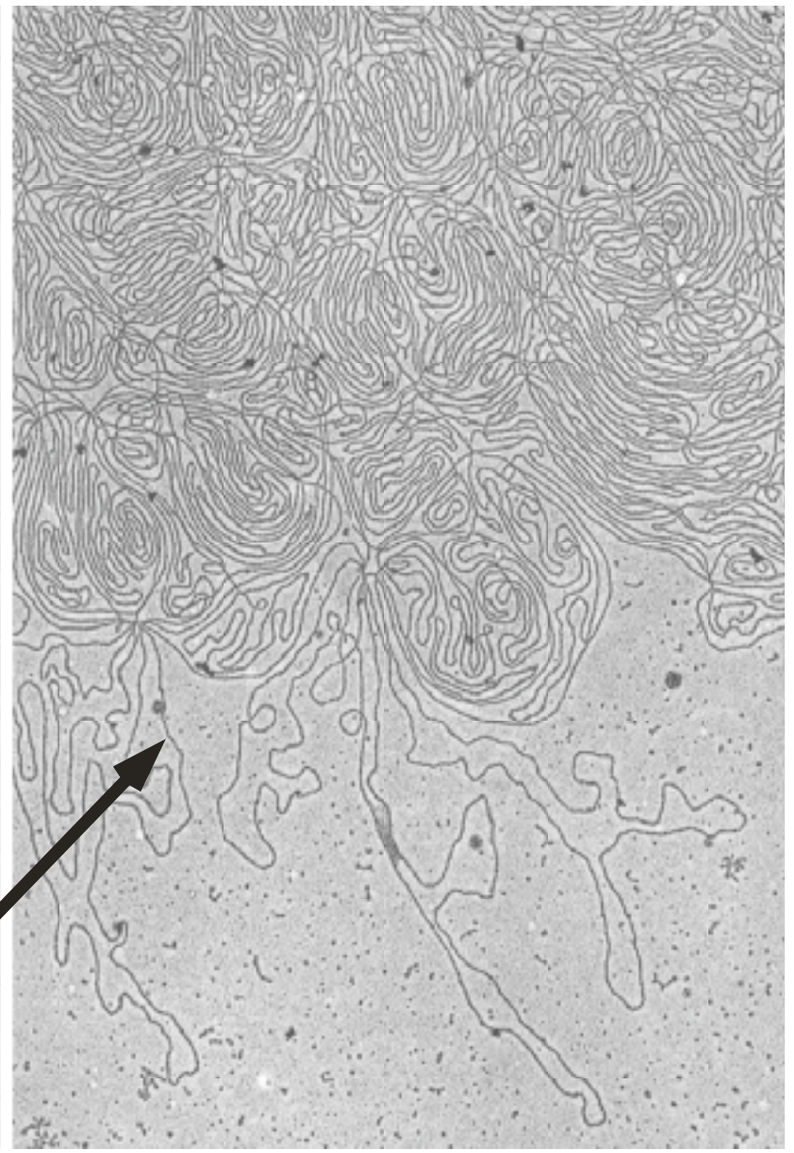
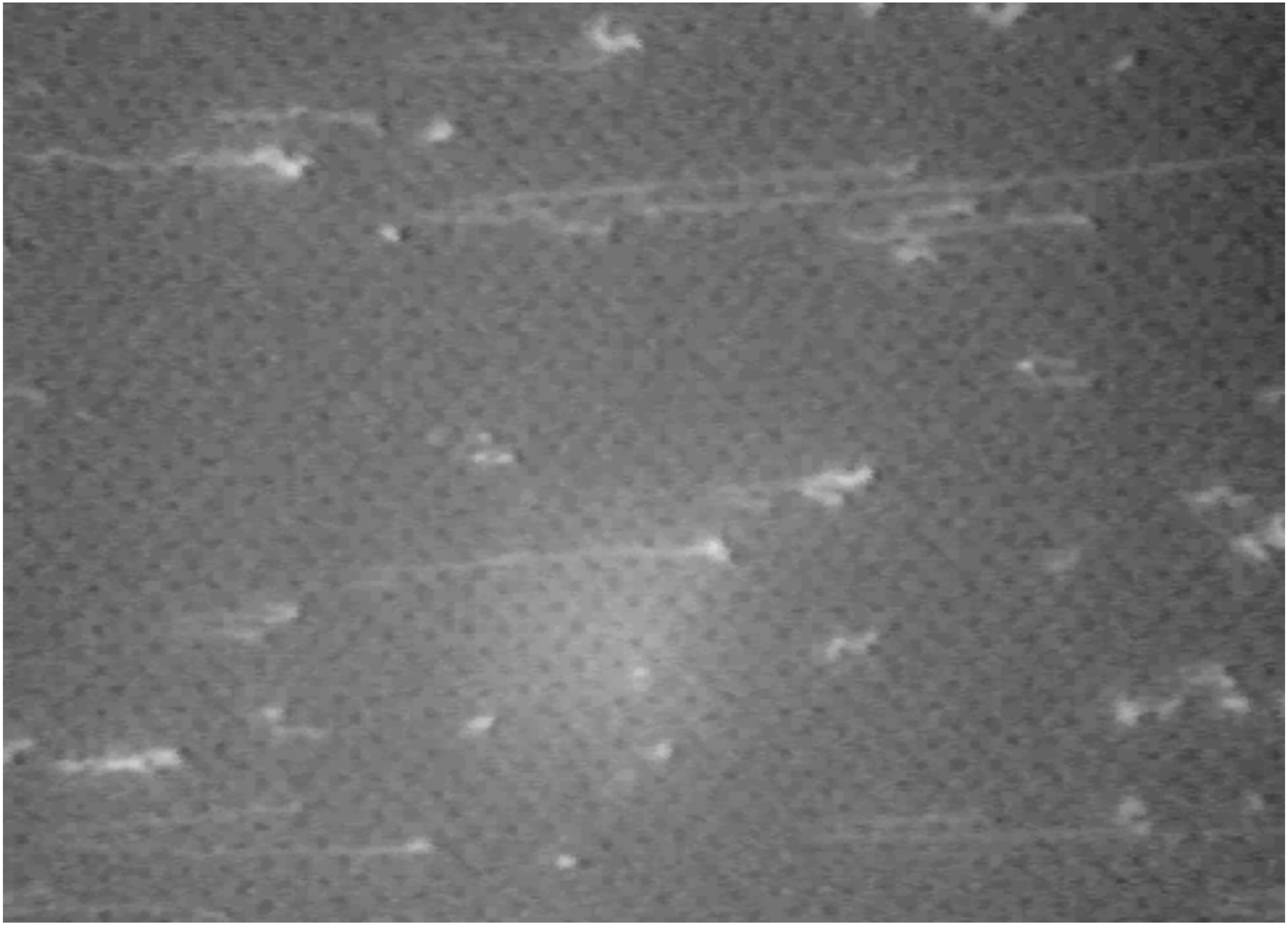


Fig. 5.2. Matching of base pairs.

How do you fit 1 meter of DNA into a 5 micron diameter ball? Very carefully!







$$Stress = (Force)/(Area) = \sigma = \frac{F}{A}$$

$$Strain = (Deformation)/(length) = \epsilon = \frac{x}{L}$$

$$Young\ Modulus = (Stress)/(Strain) = E = \frac{\sigma}{\epsilon}$$

Note: E has units of Energy/volume = Pressure.

"Soft" condensed matter physics: a "little" math reveals that for a cube of Tofu of side length L:

$$f = \frac{1}{2\pi L}[\frac{E}{\rho}]^{1/2}$$

I estimate E for Gweilo Tofu is about 10^5 ergs/cm^3

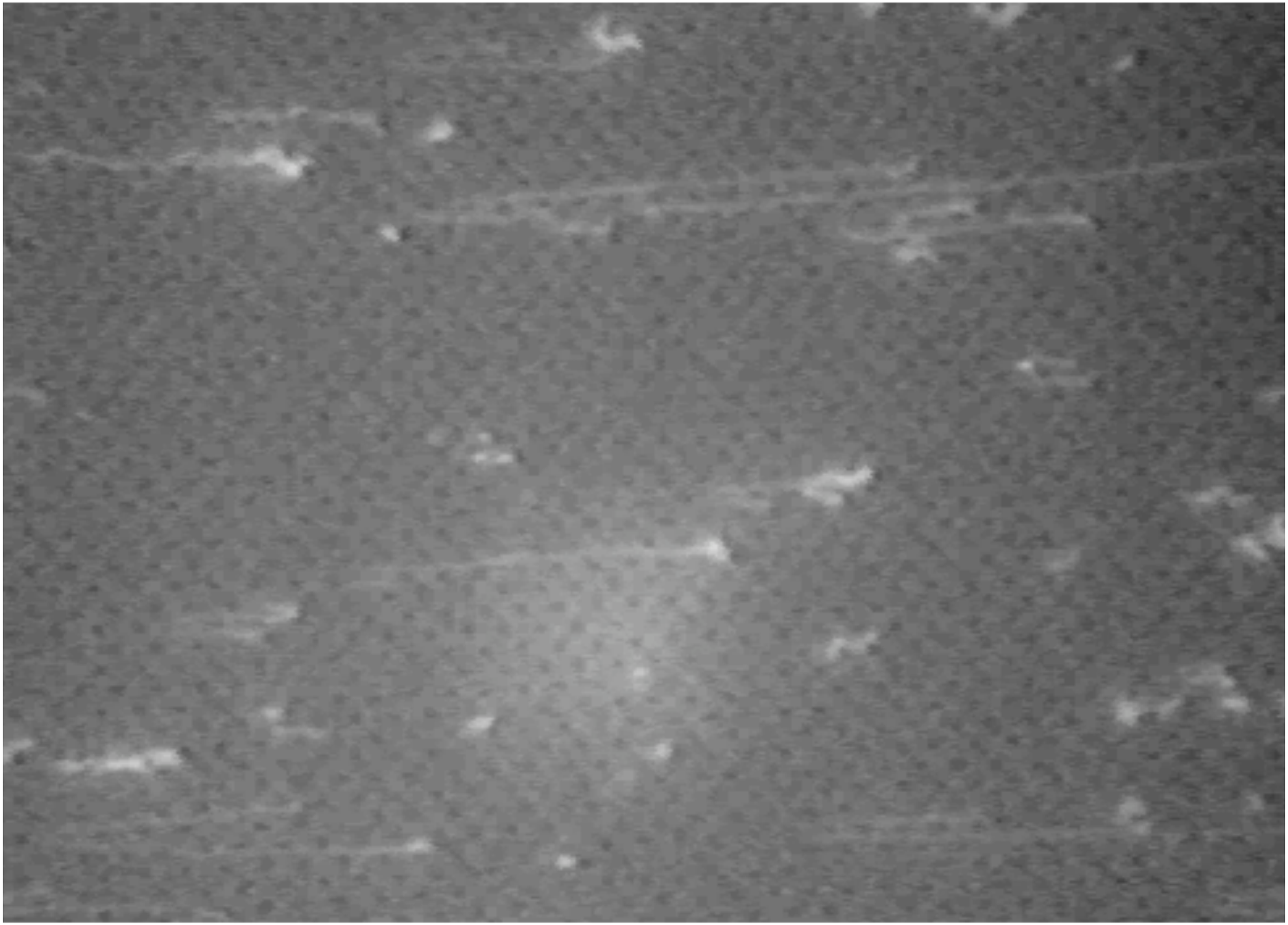
For DNA, calculate the energy to bend a rod of length L and diameter d into a curve of radius of curvature R. I'll wait while you do it.

$$U(R) = \frac{LE\pi d^4}{64R^2}$$

The persistence length p of DNA is the average radius of curvature due to thermal excitations (kT):

$$p = \frac{E\pi d^4}{64k_BT} \sim 50nm$$

E of DNA is about 10^{10} egs/cm^3, quite high.



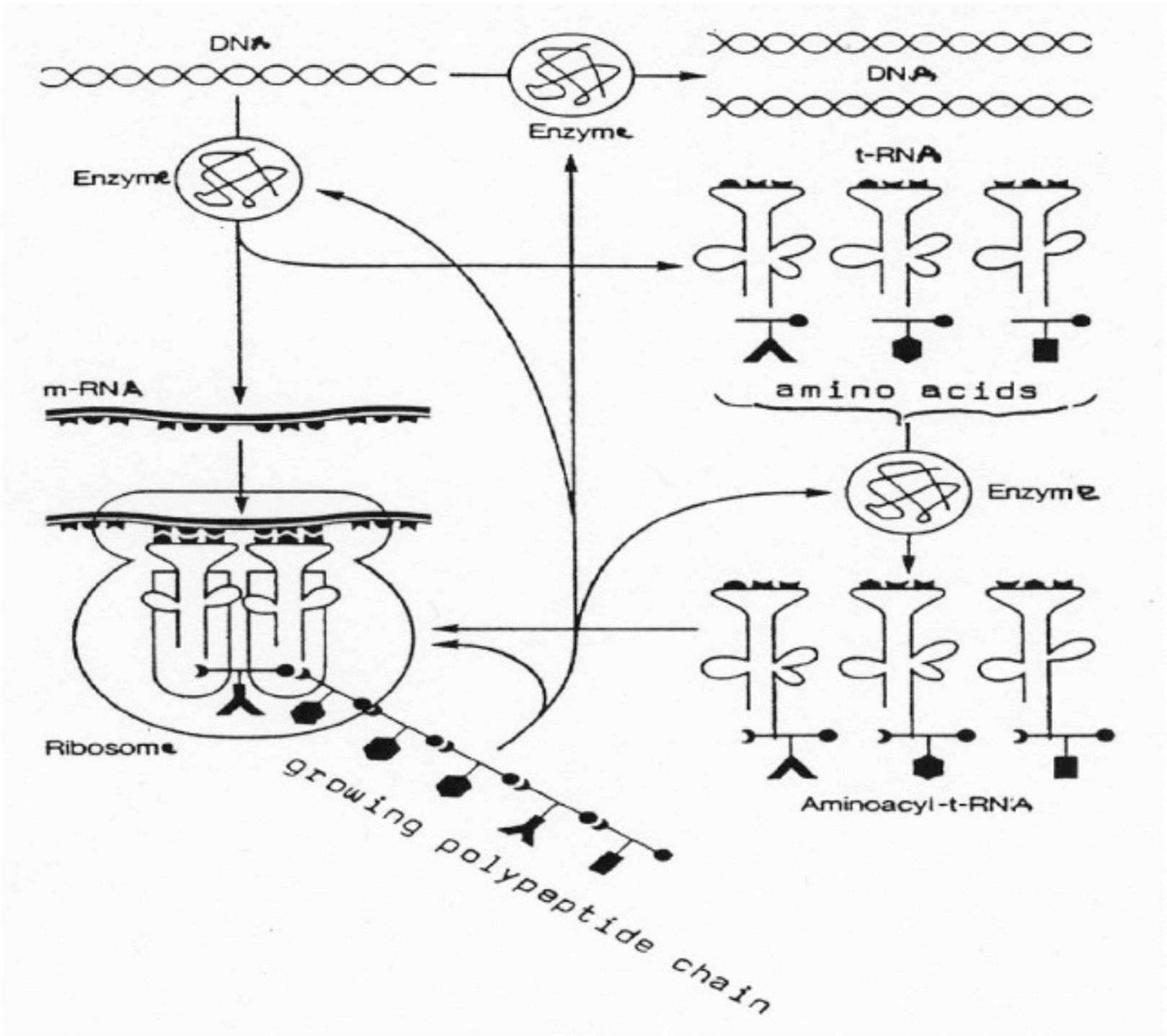
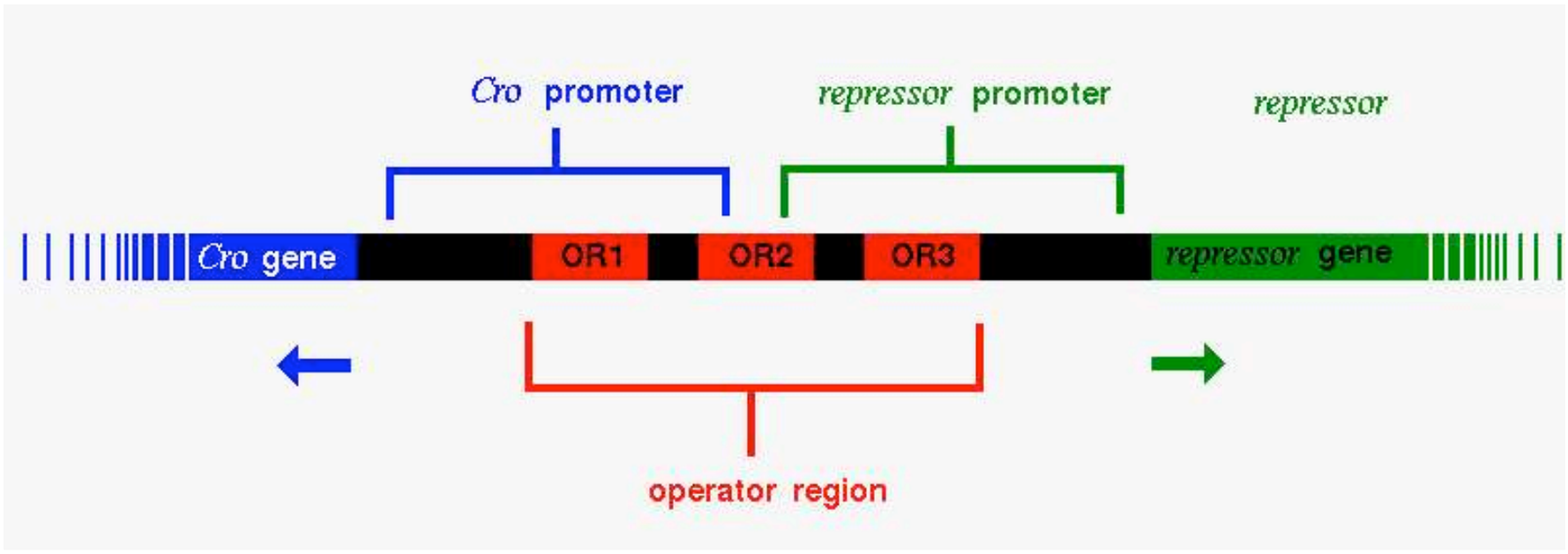


Fig. 6.1. DNA replication and direction of protein synthesis via mRNA (simplified).

Who cares?



Control of gene expression: what turns a gene on and off?

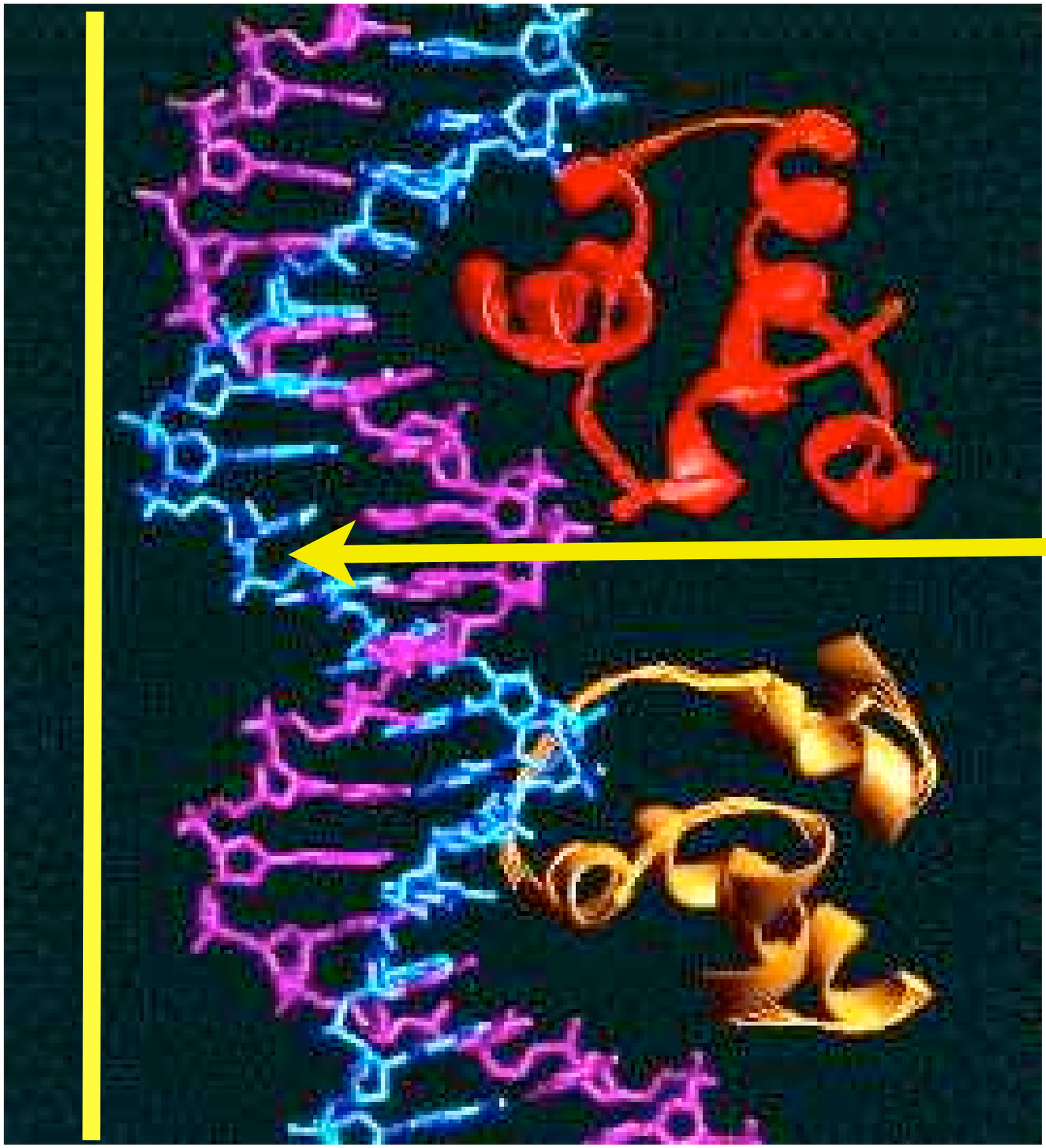


Figure 1.2: High resolution X-ray structure the 434 repressor complex. This is from the Protein Data Bank, structure PDBXXX.

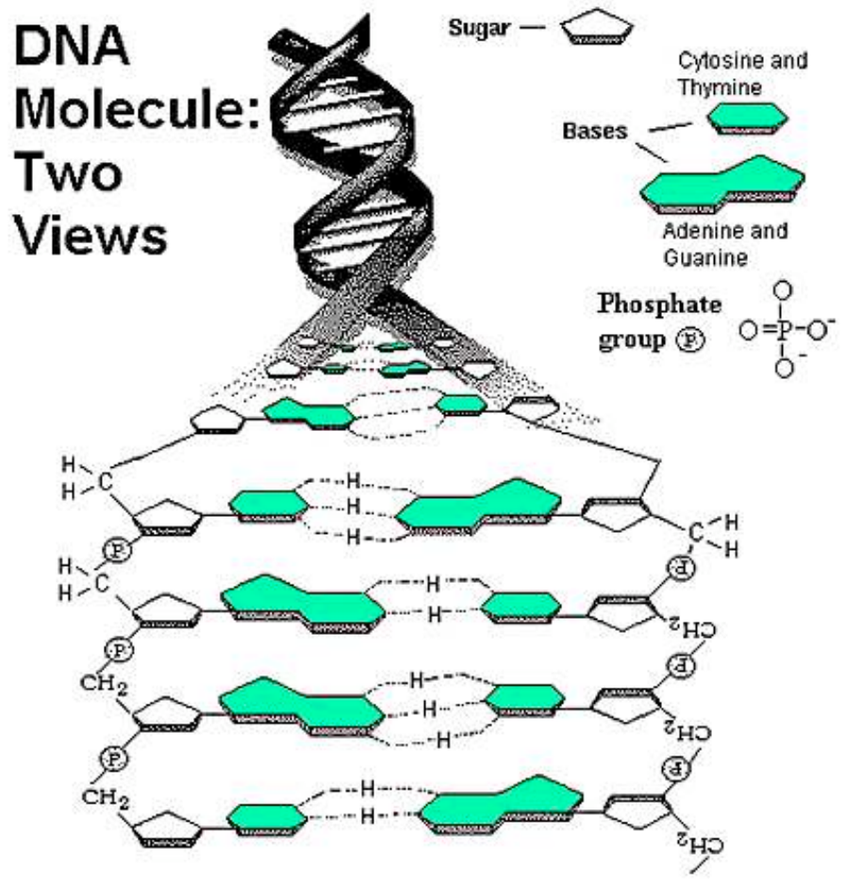


Figure 1.3: Guess what.

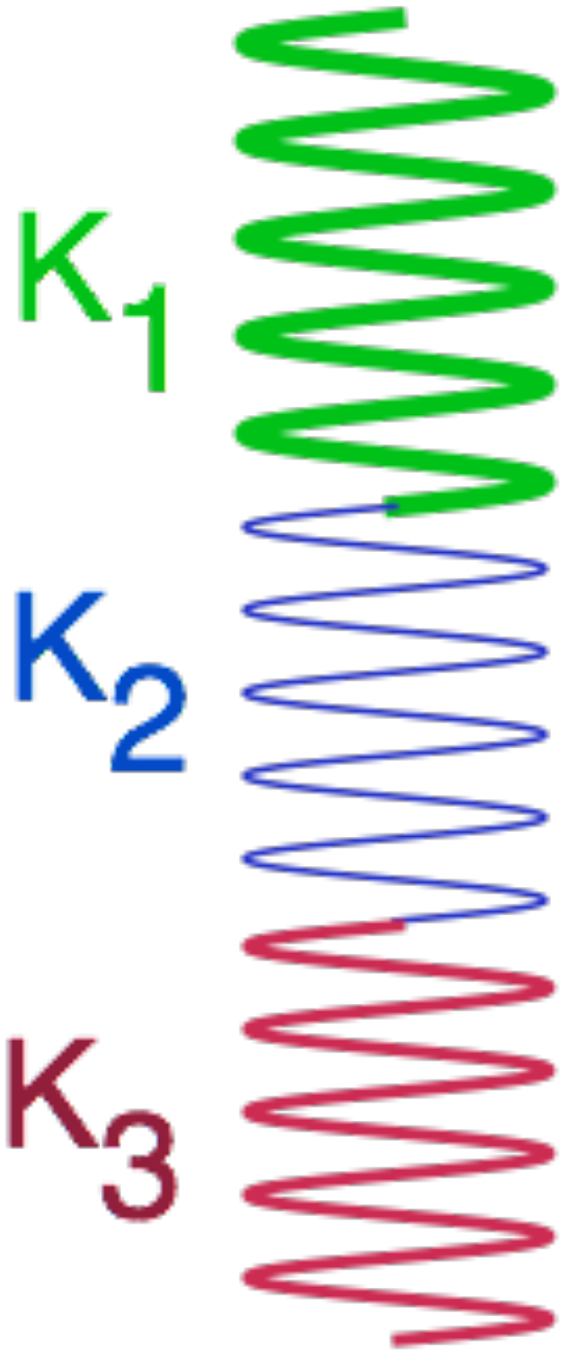
The Young Modulus E and Shear Modulus G of DNA are sequence dependent. G-C basepairs are about x5 stiffer than A-T basepairs.

I discovered this is 1983 with my buddy Mike Hogan at Princeton University, published in Nature.

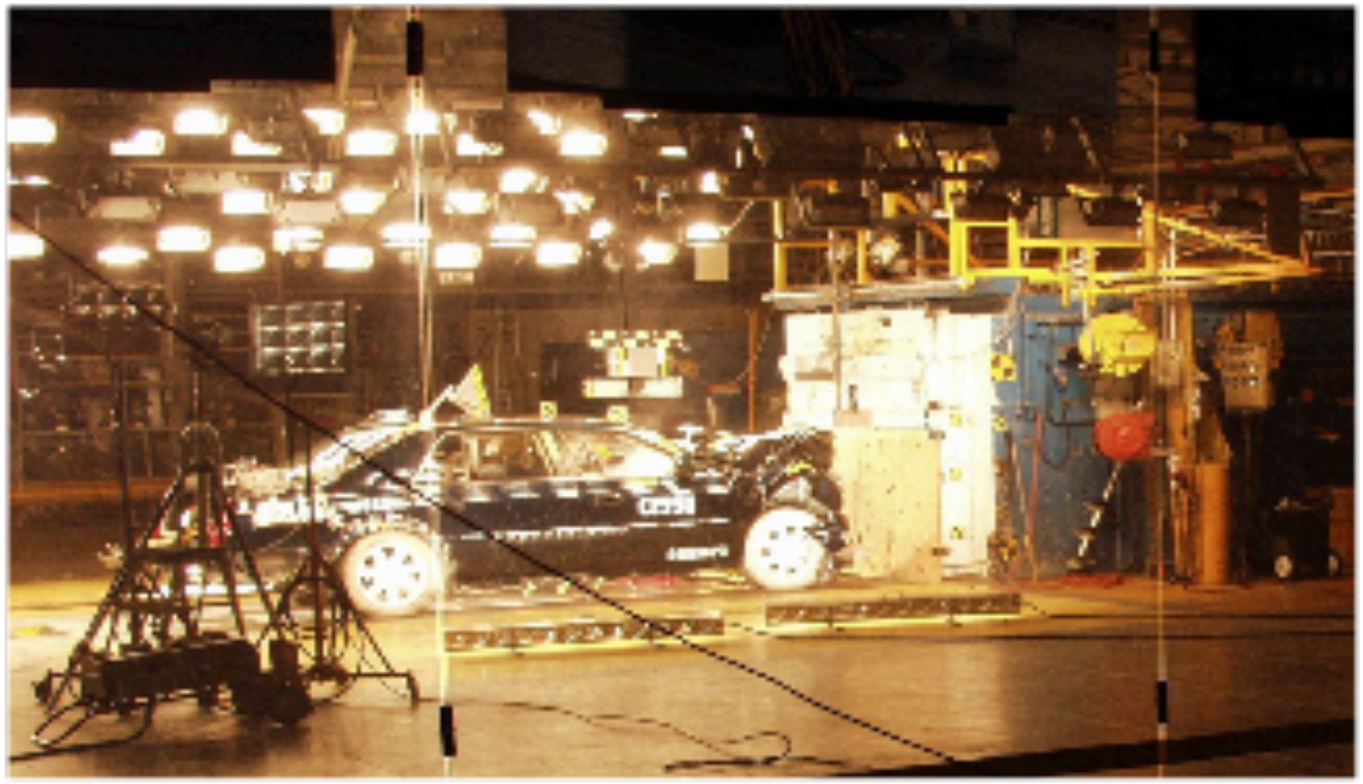
Nobody believed me or Mike because it was not a consensus opinion.

I lost all my grants, my colleague Mike Hogan didn't get tenure. I am still angry.

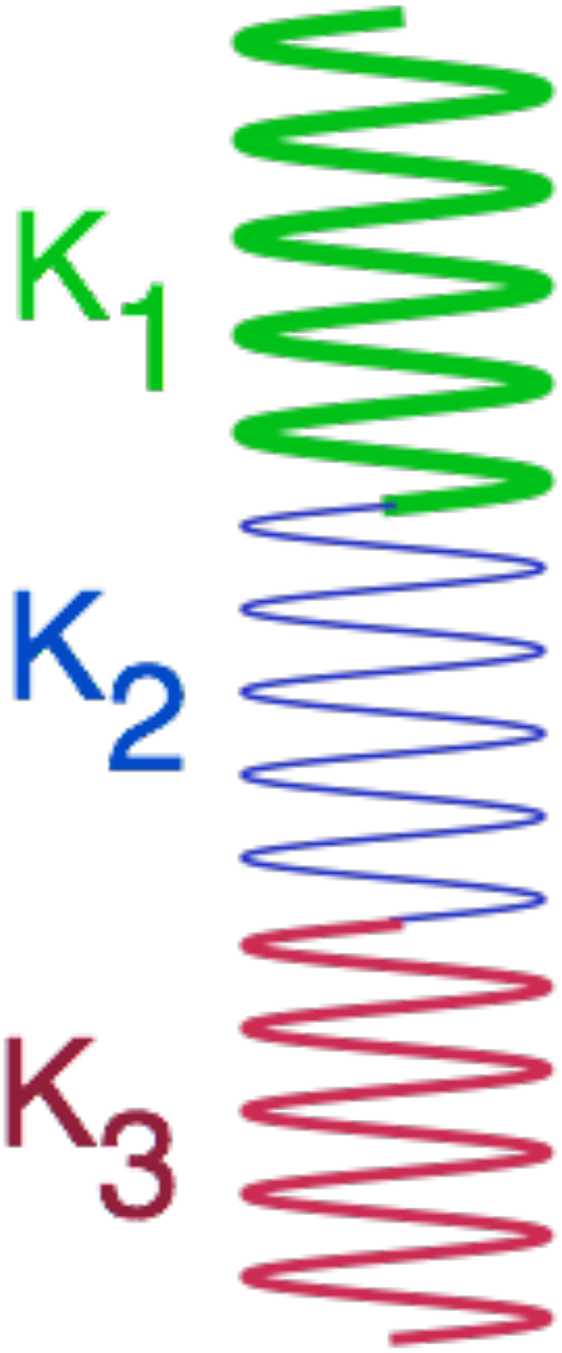
Lesson: don't do basic research unless you can stand the pain.



(note unlucky license plate number)



$$1/K_{tot} = 1/K_1 + 1/K_2 + 1/K_3$$



(note unlucky license plate number)

$$1/K_{tot} = 1/K_1 + 1/K_2 + 1/K_3$$

IF you believe my measurements, you can treat DNA a chain of little springs representing the base-pairs, and knowing the sequence and how the stiffness varies with local sequence calculate the sequence averaged bending spring constant B and twisting spring constant C for a length L:

$$\frac{1}{C} = \frac{L}{I_p}[\frac{f_{AA}}{G_{AA}} + \frac{f_{AG}}{G_{AG}} + \frac{f_{GG}}{G_{GG}}]$$

$$\frac{1}{B} = \frac{I}{I_A L}[\frac{f_{AA}}{E_{AA}} + \frac{f_{AG}}{E_{AG}} + \frac{f_{GG}}{E_{GG}}]$$

Then use this stiffness to calculate the energy stored in a twist and a bend due to the protein distorting the DNA as 434 repressor does:

$$U_{twist} = \frac{C\phi^2}{2L}; U_{bend} = \frac{B}{2R^2}$$

Finally, use the Boltzmann relation to see how binding constants K vary with the energies required to bend and distort the DNA/protein complex:

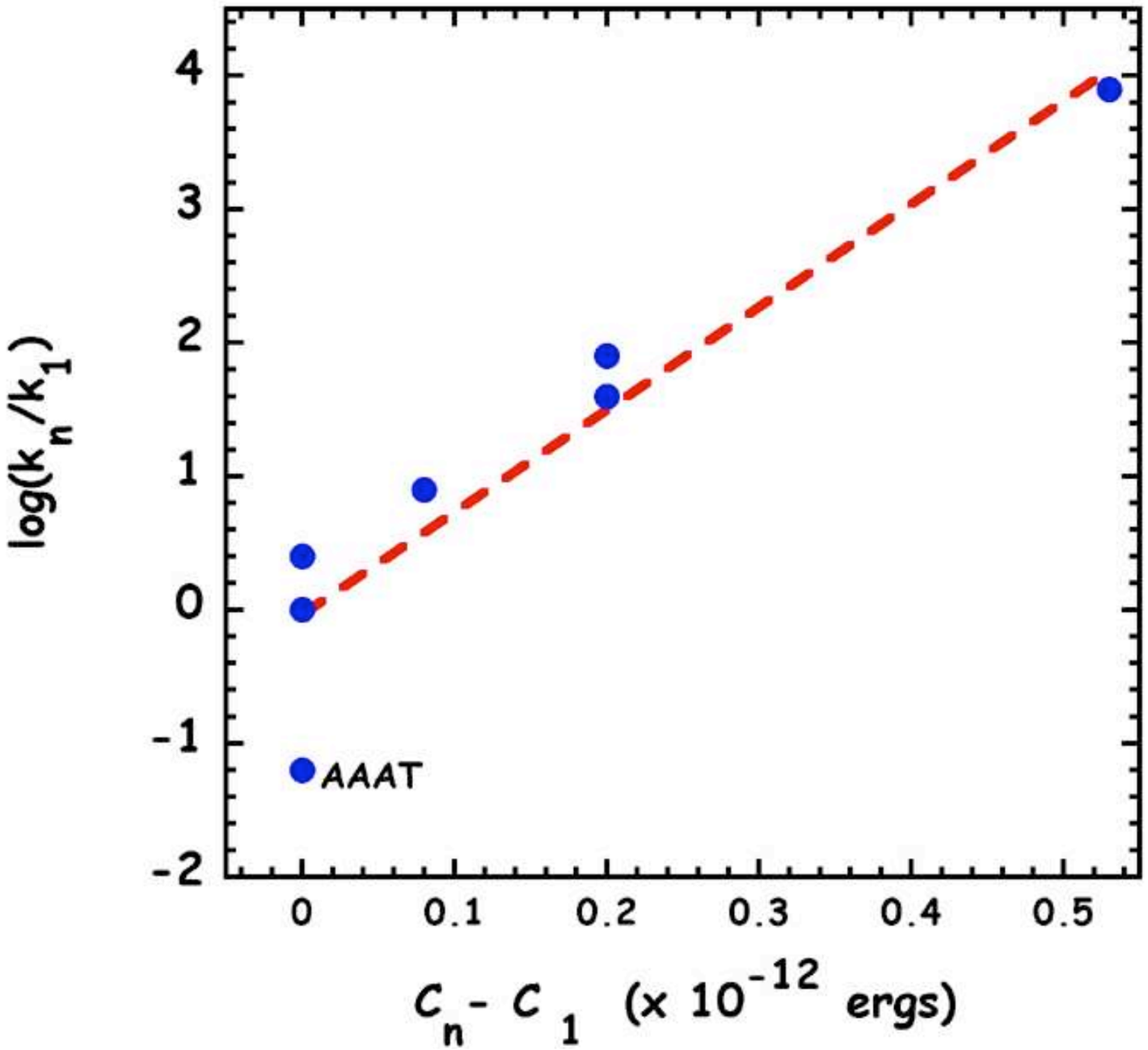
$$K(U) \sim K_o \exp[-U/k_B T]$$

$$\log[\frac{K_1}{K_2}] \sim \frac{U_2 - U_1}{k_B T}$$

Repressor Binding Coefficients

Name	Sequence	f_{AA}	f_{AG}	f_{GG}	C	B	$(K_1/K_n$	$\log(K_1$
1	ATAT	1	0	0	0.37	1.15	1	0
2	TTAA	1	0	0	0.37	1.15	1.5	0.4
3	CTAG	1/3	2/3	0	0.57	1.8	5	1.6
4	GTAC	1	0	0	0.57	1.8	5	1.6
5	GTAT	1	0	0	0.45	1.4	2.5	0.9
6	AATT	1	0	0	0.37	1.15	1	0
7	AAAT	1	0	0	0.37	1.15	0.3	-1.2
8	ACGT	1	0	0	0.9	2.8	50	3.9
9	AGCT	1	0	0	0.9	2.8	50	3.9
10	AGAT	1	0	0	0.57	1.8	7	1.9

The sequence of the distorted DNA in the 434 repressor-DNA complex is known, and you can predict how binding constants should depend on sequence.


4
3
2
1
0
-1
-2
AAAT
0
0.1
0.2
0.3
0.4
0.5
log(k_n/k_1)
C_n - C_1 (x 10^-12 ergs)

3) Peering into the future.

1. A woman who seeks to be equal with men lacks ambition.

2. The greatest danger today could be your stupidity.

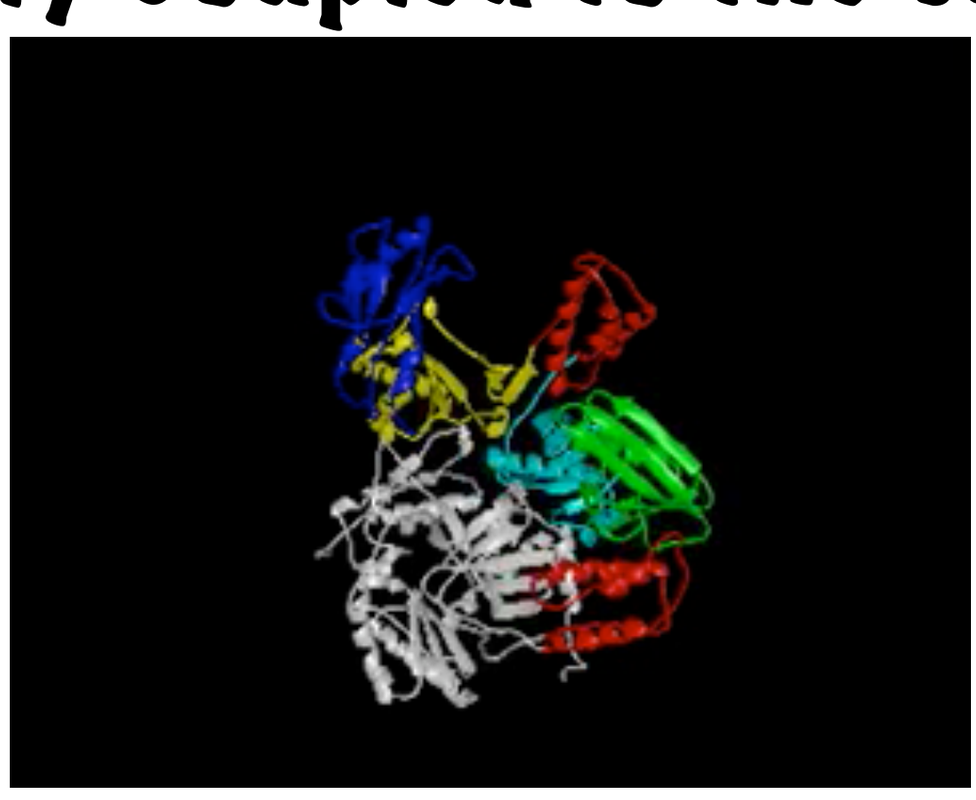
3. He who laughs last is laughing at you.

4. It's over your head now. Time to get some professional help.

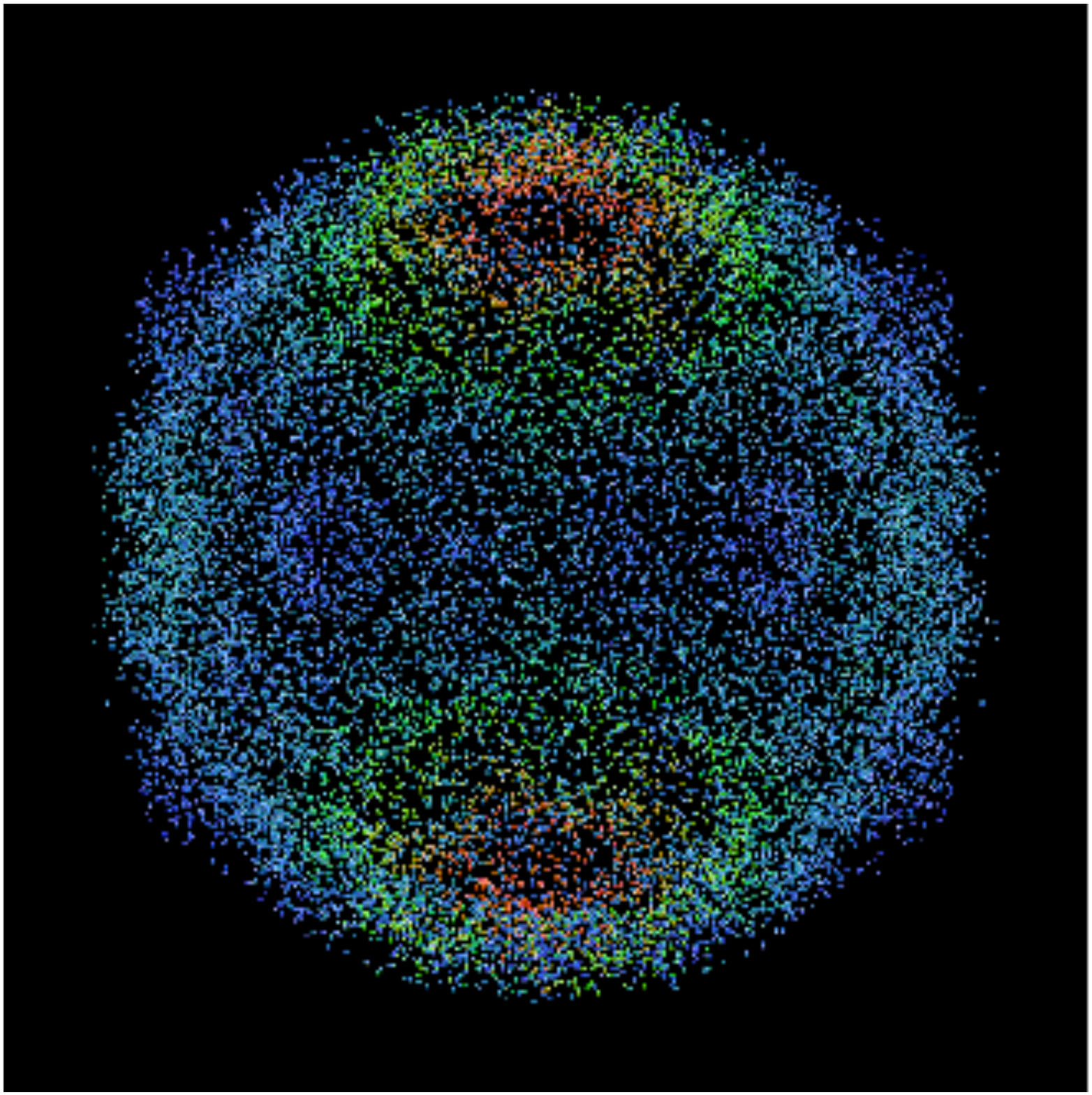
I don't remember reading of this in The Tao

1. Will we ever understand the basic physics of molecular motors?

I think the answers lie in the large-scale collective modes of the protein, out where the mode energies are roughly k_BT (terahertz), and strongly coupled to the solvent.

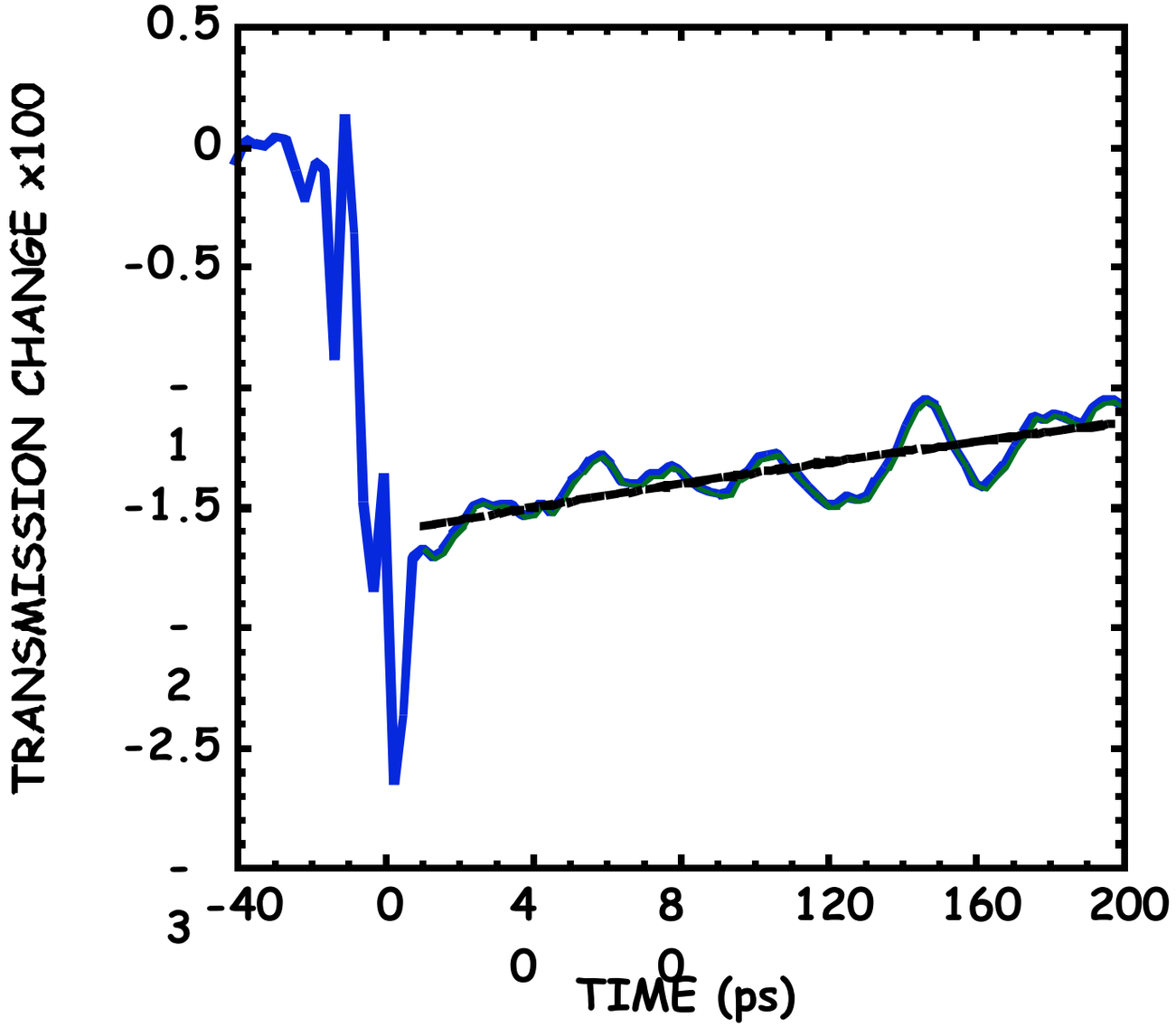


"Collective dynamics of HIV-1 reverse transcriptase: Examination of Flexibility and Enzyme Function"
I. Bahar, B. Erman, R. L. Jernigan, A. R. Atilgan, & D. Covell J. Mol. Biol. 285, 1023-1037, (1999)



Calculation of an entire virus capsid [2]. The virus capsid is made up of hundreds of proteins that form the protective casing for a nasty piece
viral DNA. The virus capsid is known to exist in an immature form with a given symmetry. On maturation, the capsid changes into a differen
symmetry.

Here is the experimental result for FIR pump/probe on bacteriorhopsin at 87 microns (115 cm^{-1}). Do you believe it? It is a very strange result. Lifetime of 500 ps= Q of about 300. You should be worried.



But I quit, I've run out of time and money.

2) Will we ever understand how proteins bind with high basepair selectivity to DNA at some basic level of physics?

DNA bending is an important structural feature for indirect readout in protein-DNA recognition……We applied a new all-atom Monte Carlo (MC) algorithm that combines effective sampling with fast conformational equilibration. The resulting MC ensembles resemble the corresponding high-resolution crystal structures very well…..Distinct bending is observed for the E2-DNA binding site with a central AATT linker in contrast to an essentially straight DNA with a central ACGT linker…. Contributions of specific base pair steps to the overall bending are shown in terms of local structural parameters. The analysis of conformational substates provides new insights into the energetic origins of intrinsic DNA bending.

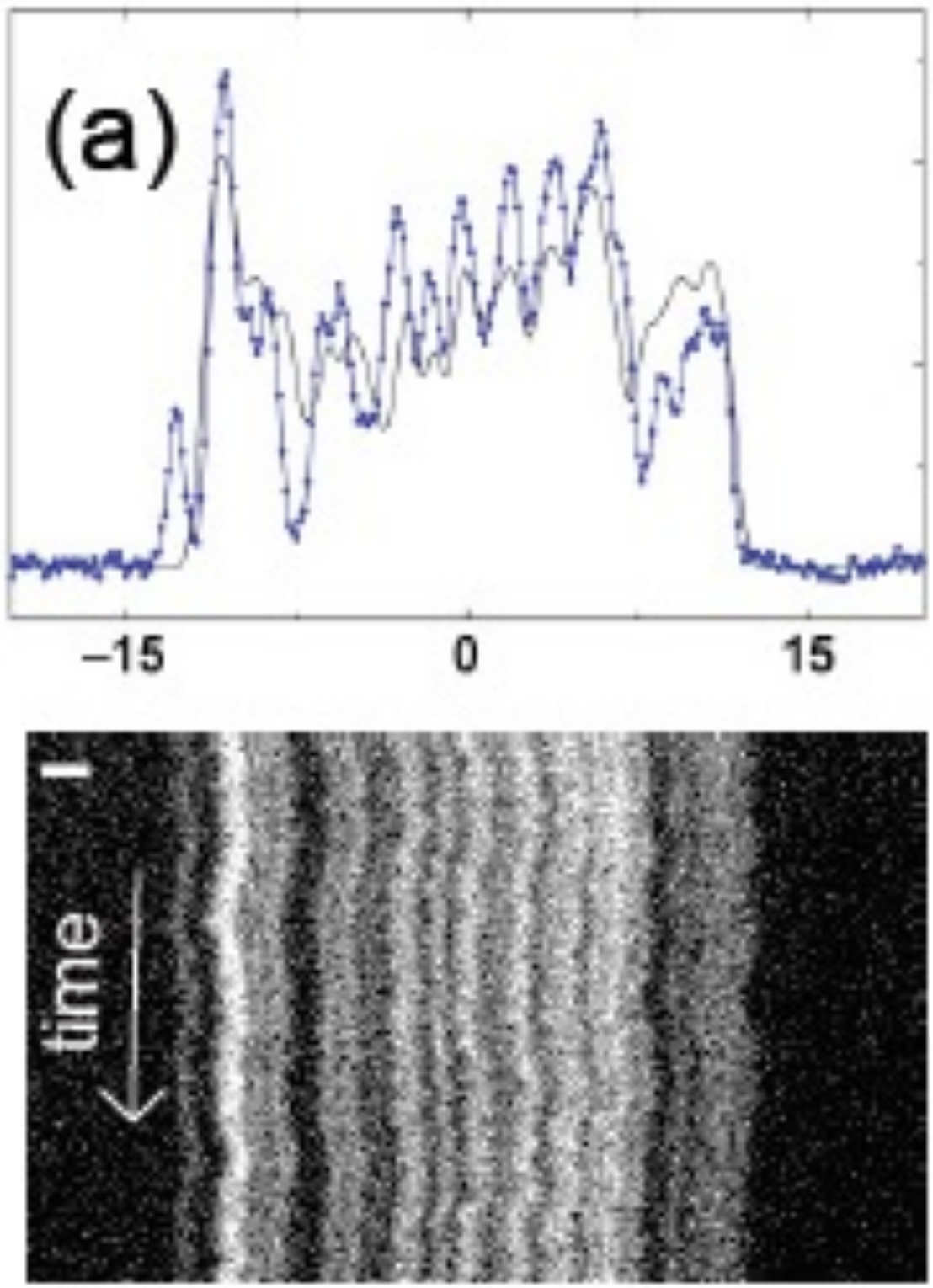
Structural and Energetic Origins of Sequence-Specific DNA Bending: Monte Carlo Simulations of Papillomavirus E2-DNA Binding Sites
Structure, Volume 13, Issue 10, Pages 1499-1509, 2009
R. Rohs, H. Sklenar, Z. Shakked

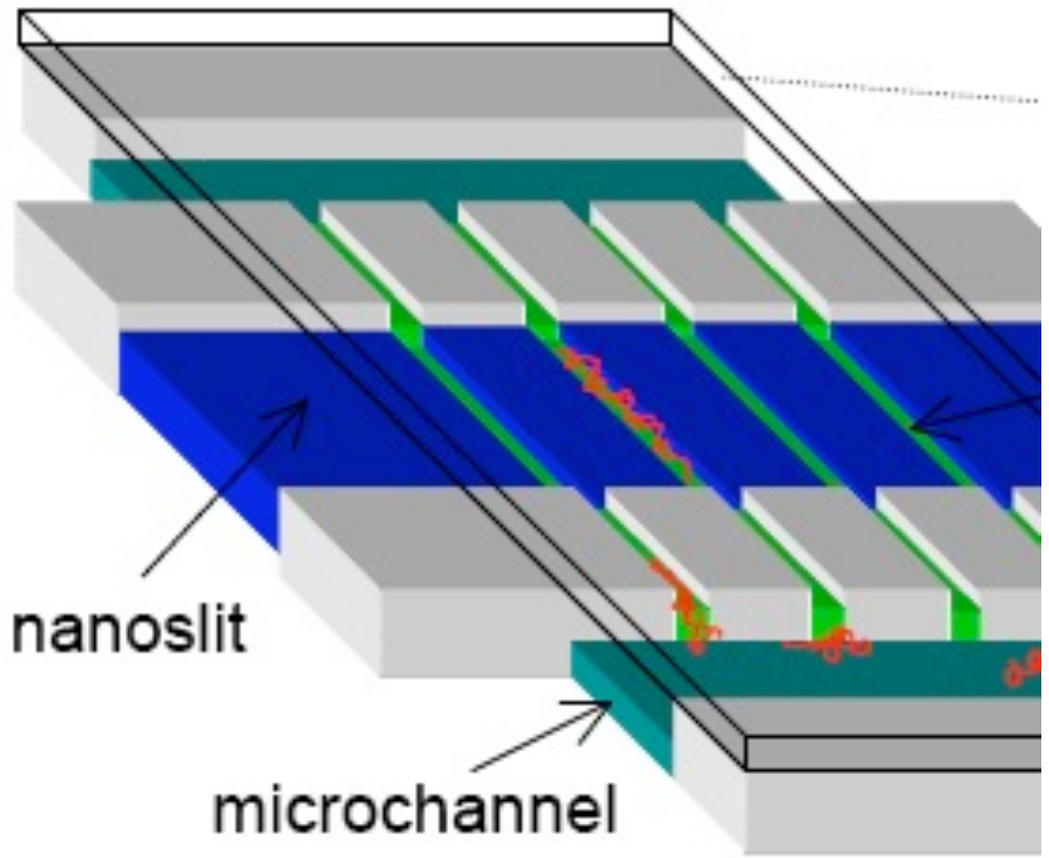
I'm not sure people are interested in analytical analysis anymore, if the computer says so then that is the end of the story.

But, to anticipate TUNG Chih-kuan a bit, my former student Walter Reisner has been playing with melting DNA in nanochannels.

Note that thermodynamically the weaker the bonds, the lower the melting point.

It is well known that DNA melting points are sequence dependent, I wonder if you can visualize that and get the sequence of DNA....





Oh yes!

THANKS!