INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
I.C.T.P., P.O. BOX 586, 34100 TRIESTE, ITALY, CABLE: CENTRATOM TRIESTE

H4.SMR/642 - 11

College on Methods and Experimental Techniques in Biophysics

28 September - 23 October 1992

Protein Folding
(Transparencies)

J.N. ONUCHIC
University of California, U.S.A.

These are preliminary lecture notes, intended only for distribution to participants.

Protein Folding - Lecture 1

"PROTEINS"

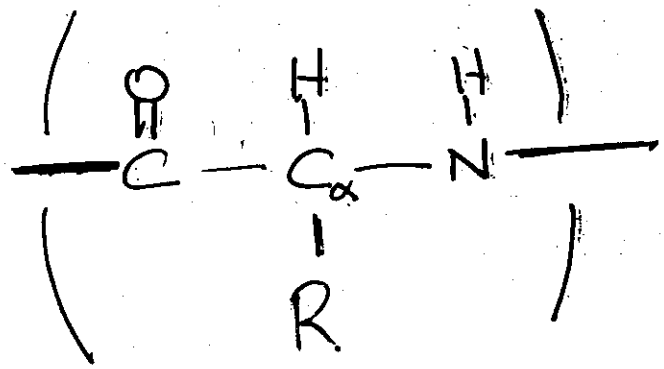
MAIN TOPIC OF STUDY DURING THIS COLLEGE

- Several techniques will be described - Most of the applications to proteins

WHAT ARE PROTEINS?

Proteins are heteropolymers of Amino-Acids

AMINO ACID



R characterizes the Amino-Acid

A protein can be described by its sequence of Amino-Acids !!!

Looks obvious !!!

But we will elaborate on this statement during our lectures.

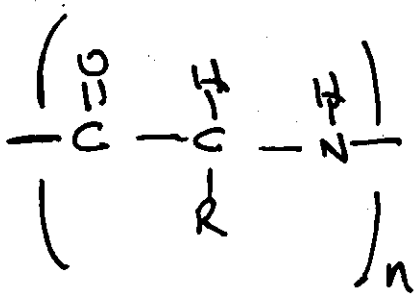
Why we care about proteins?

Biological Function — Proteins do the work!

For example:

- Energy Conversion
- Oxygen and CO_2 transport
- Catalysis (Enzymes)
- Immune Recognition
- Information Process
- ....

However — Function (and dynamics) depends on the 3-D structure of the protein — Not the sequence



——D proper conditions

UNIQUE STATE
(FOLDED STATE)

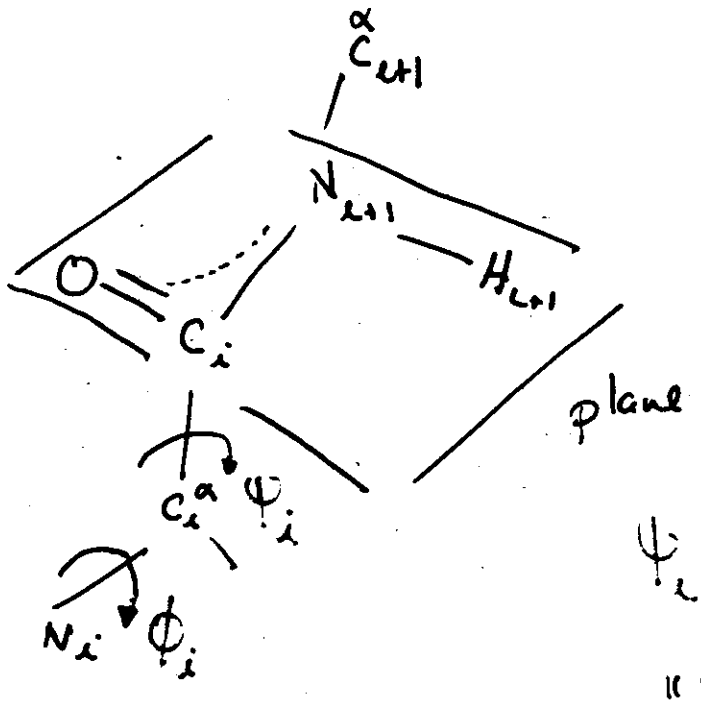
ENERGETICS AND GEOMETRY OF HOW PROTEINS ORGANIZE THEMSELVES IN 3D

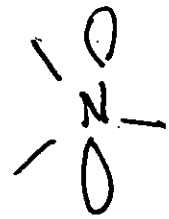
- Lubert Stryer, "Biochemistry", chapter 2, W.H. Freeman, New York, third edition 1988
- T.E. Creighton, "Proteins: Structures and Molecular Properties", W.H. Freeman, New York 1984

(People without a reasonable knowledge on this subject should look for me.

Highlight 3 points

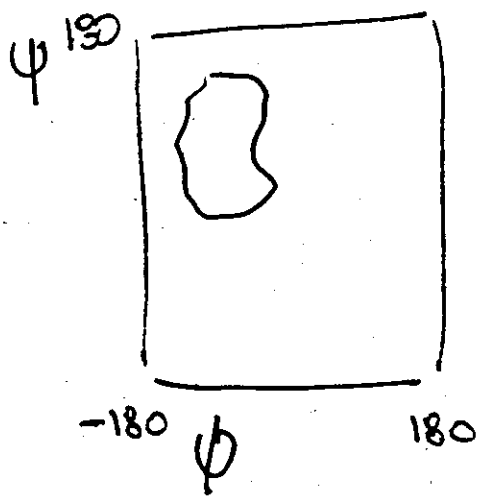
(1)





Ψ_i, ϕ_i define the "local" 3-D structure of the protein

Rama Chandra Plot



Possible values of

ψ and φ

"Geometrical Constraints"

→ Possible local geometries

---

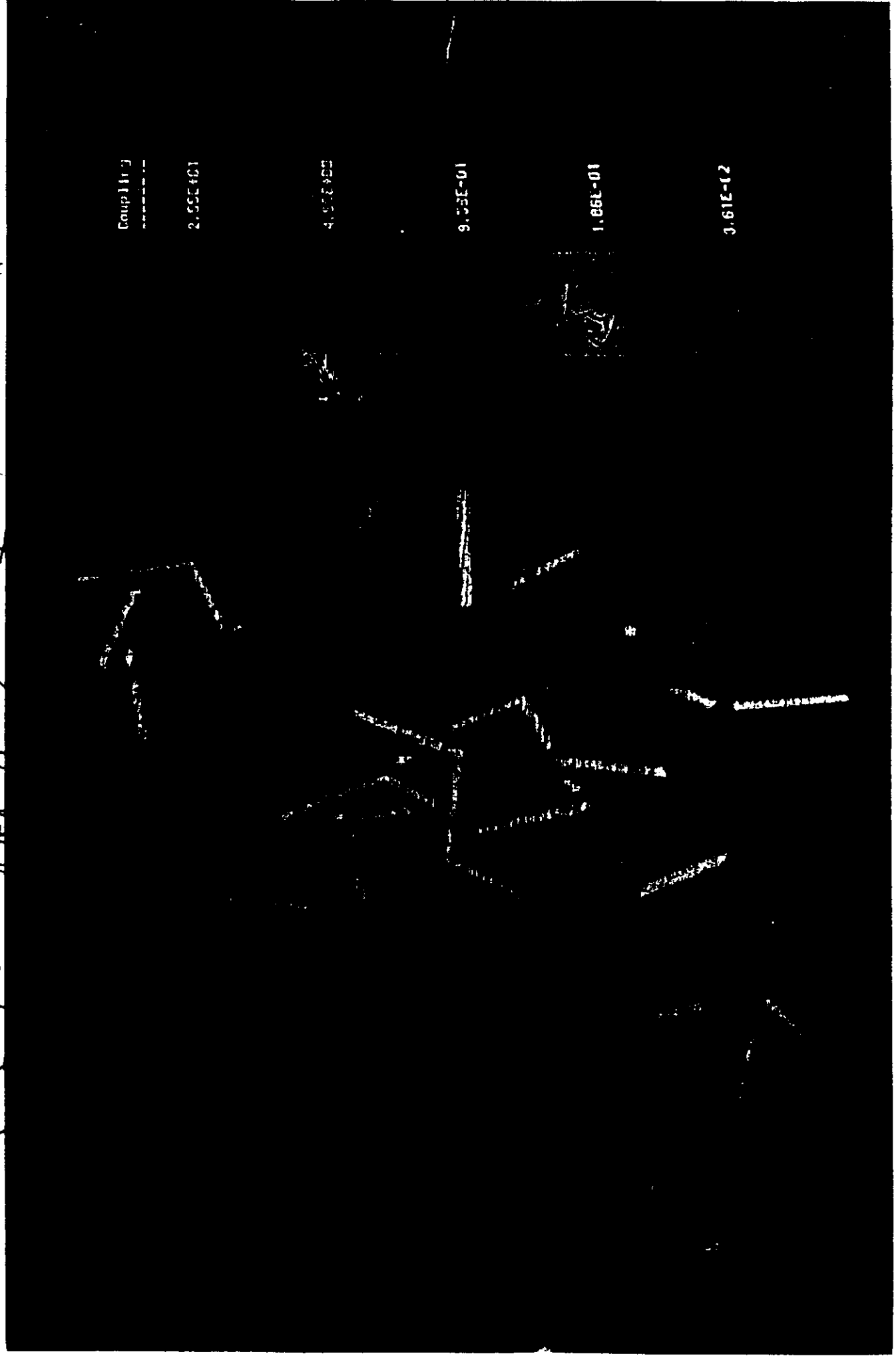
(2) Hydrogen Bonds $\overset{O}{\overset{\|}{C}}$ and $\overset{H}{N}$ of AA's like to H.B.

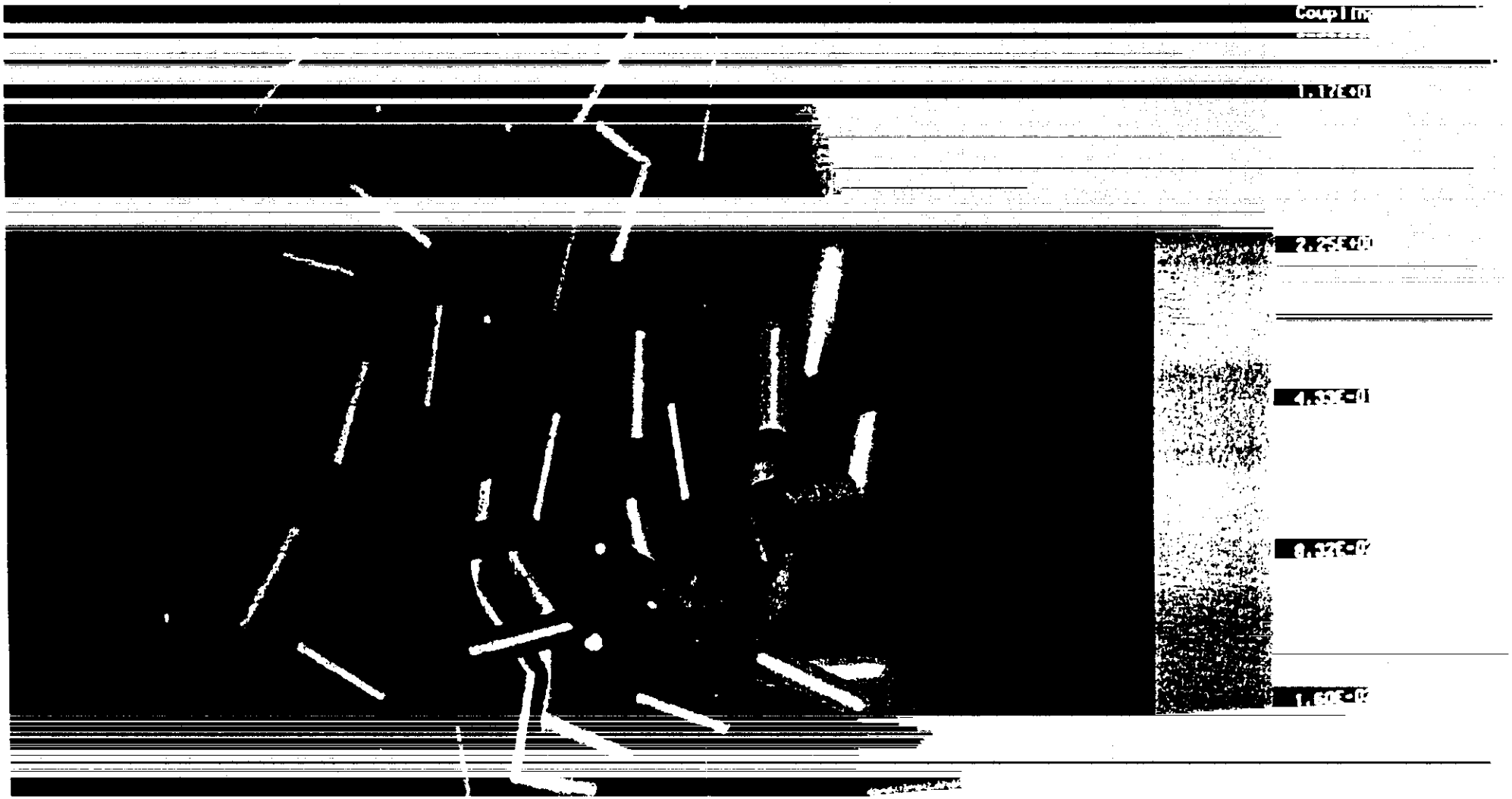
→ Two possible periodic structures (ϕ_i, ψ_i) are periodic

- α - Helix
- β - Sheet

---

(3) Solvent Effects - hydrophobic interaction


Coupling
2.55E+01
9.53E-01
1.86E-01
3.61E-02


Coup l ing
1.12E+01

Primary Structure – sequence

Secondary Structure – local geometry (α helix, β sheet)

tertiary Structure – global geometry

(quaternary structure – more than 1 sequence)

Look the importance of the effects described before

SEQUENCE-STRUCTURE RELATIONSHIP

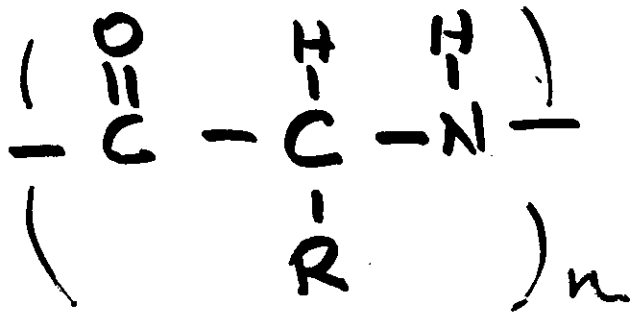
However, for Biological Proteins, one sequence defines one structure

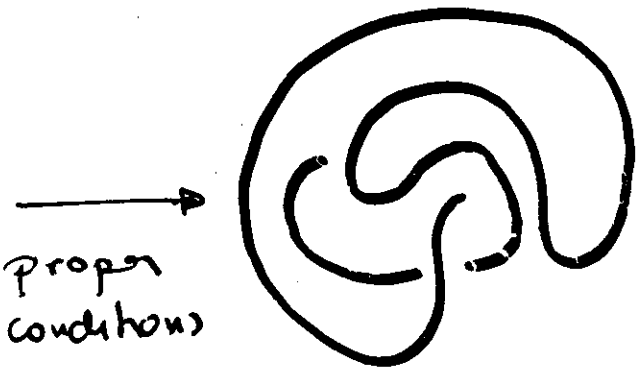
Second Genetic Code

"All the information is on the sequence"

Proteins

Amino Acid Polymer





R - defines the AA.

UNIQUE STATE (FOLDED STATE)

NOT ALL SEQUENCES FOLD !!!

Protein Folding Problem

CAN YOU PREDICT THE 3-D STRUCTURE BASED ON THE AMINO ACID SEQUENCE ?

CAN YOU PREDICT IF A STRUCTURE WILL FOLD ?

FOLDING MECHANISM

(4) (4)

Anfinsen '60's (Ribonuclease)

Ribo
inac

SMALL PROTEINS (GLOBULAR) FOLD (REFOLD)
IN THE ABSENCE OF ANY CATALYTIC MOLECULE

- **THERMODYNAMIC HYPOTHESIS**
  (GLOBAL STABLE MINIMUM)

.evinthal '68

TOO MANY STATES TO PROBE TO FIND THE RIGHT ONE !!! **PARADOX**

each AA. 2 positions 2^N states $N = 100$ 10^{30}

PROTEIN PROBE 10^{12} **states/sec** 10^{18} s

To save the T.H.

Baldwin '83 → Folding Pathway

FACT Not All Proteins fold only biological ones

∴

Not all sequences meet the thermodynamic and kinetic criteria

---

Theoretical Efforts → **MECHANISM?**

Analytical
- Mean Field
  - Spin Glass
  - Polymer
- Mean Field
  - thermodynamics
    - folding time → - Dill
      - Shakhnovich
      - Bryngelson and Wolynes

Lattice Models
- Preserved Sequence information
  - sequence – energy – structure
    - exact enumeration
  - Lattice kinetic simulations
    - Folding Pathways ??

Dill, Shakhnovich, Skolnick, ...

THERMODYNAMICS

① Folded State Unique

- Two state system

- First Order Phase Transition

ORIGIN ?

- If this is correct - we need some energy gap between the lowest state (folded) and the continuum of other (non-folded) states

DYNAMICS - How to beat Levinthal?

- Diffusion-Collision Model

  FIRST SECONDARY STRUCTURES
  SECOND TERTIARY STRUCTURE

- Collapse-Reconfiguration Model

  FIRST COLLAPSE -
  SECOND Boundary Structure

- Framework Model

  single secondary structure
  nucleates the process

→ This dynamical views influence even the

Results

- Folded Structures – stable (thermodynamical stability)

  ??

- Principle of Minimal Frustration

  Band W HOW ?

- Lattice kinetics – Pathways

- Potentials Seq. – St. relationship

  N- Letter Code