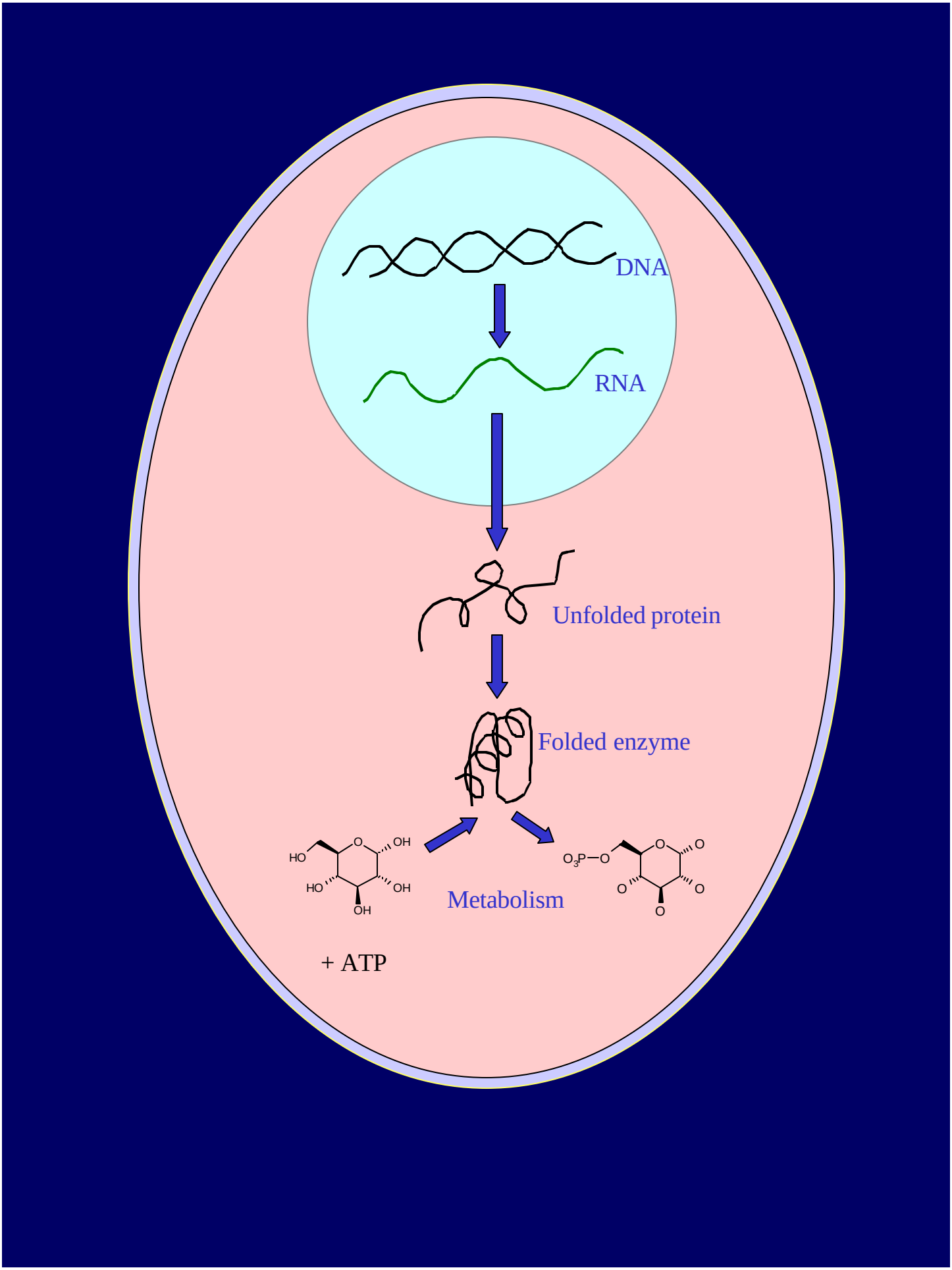
Predicting Physico-Chemical Properties of Gene Products

Michael K. Gilson, M.D., Ph.D.
Associate Professor
Center for Advanced Research in Biotechnology
University of Maryland Biotechnology Institute

gilsonlab.umbi.umd.edu

Molecules, Properties & Applications


DNA
RNA
Unfolded protein
Folded enzyme
HO
OH
HO
OH
OH
O
Metabolism
O_3P—O
O
O
O
O
+ ATP

RNA

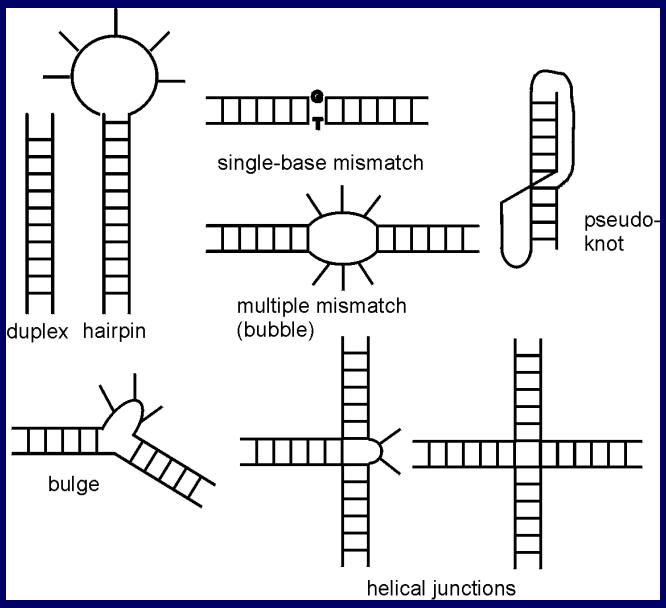
Secondary structure: H-bonding

Tertiary structure: 3D

Binding affinities

- Other RNA structures; e.g., antisense)
- Proteins; e.g., Tat/Tar
- Ions; e.g., Mg^{2+}

Catalytic properties (ribozymes)



http://pps98.cryst.bbk.ac.uk/assignment/projects/wieczorek/Dominika/page4.htm

Proteins

Secondary structure: β-sheet, α-helix

Tertiary structure: 3D

Quaternary structure: assemblages

Binding affinities
- Proteins
- DNA, RNA
- Small molecules; e.g., transmitters

Ionization states (pKas)

Stabilities; e.g., heat, acid

Catalytic properties: Chemistry, specificity

Spectroscopy; e.g., chlorophyll

Allostery

Energy transduction
- Chemical → mechanical
- Mechanical → chemical

“Small Molecules”

(e.g., metabolites, transmitters, membrane lipids)

Conformational distributions

Ionization states (pKas)

Biological compartmentalization

- Solubility
- Membrane permeance

Binding affinities

- Proteins (enzymes, receptors)
- DNA, RNA

Properties of assemblages; e.g.,
lipid bilayers:

- Phase transitions
- Permeabilities
- Fusion mechanisms

Applications

Protein engineering, possibly as part of metabolic engineering

Enzyme inhibitors to control metabolic flow

Designed chemical separations for metabolic products, biomass

Drug-design: small molecules and biopharmaceuticals

Gene annotation

Prediction Methods

Physics-based

Knowledge-based

Physics-Based Methods

Classical statistical thermodynamics -- evaluation of integrals over many conformations

Quantum chemistry -- solving Schrodinger's equation for a few conformations

Classical Modeling

Empirical force fields; e.g., CHARMM, AMBER, OPLS, MMx

Implicit solvent models; e.g. Poisson-Boltzmann electrostatics, generalized Born

Sampling techniques: molecular dynamics, Monte Carlo, Predominant States, global optimization

Typical Applications

- Predicting conformations of small molecules or parts of proteins
- Modeling dynamics of proteins, assemblages, around a conformation
- Modeling membranes
- Calculation of noncovalent binding affinities

Quantum Chemistry

Ab initio and semi-empirical

Analyses one or a few molecular conformations

Typical Applications

- Modeling reactivity and chemical catalysis
- Predicting and interpreting spectroscopy
- Developing empirical force-fields
- Predicting local structure (e.g., metalloproteins)

Knowledge-Based Methods

Based upon molecular similarity & databases

- Macromolecular structure (PDB, NDB...)
- Functional annotations of genes
- Structure-activity relationships of small molecules

Typical applications

- Protein & RNA structure prediction; e.g, homology,threading, substructure libraries
- Prediction of catalytic activity; e.g., via homology modeling and 3D motif detection

Directions

Models and Algorithms

Enhanced empirical force fields

- Electronic polarization
- Broader range of validity
- Increased accuracy

New serial and parallel algorithms

- classical
- quantum

Combined quantum/classical modeling

Broader application of knowledge-based methods; e.g., macromolecular binding

Membrane proteins

Validation

Data sets and databases

Prediction versus postdiction; e.g., CASP

Standards-setting

Software

Infrastructure of basic algorithms; e.g., force-field implementations and molecular manipulations

User-interfaces, especially for non-specialists and students:

- Desktop applications
- Web-based applications