Protein Engineering Past & Future:
A Personal Perspective

Ray Salemme

Levy Group
Temple University Chemistry Dept.
March 28 2018

Protein Engineering: A Personal Perspective

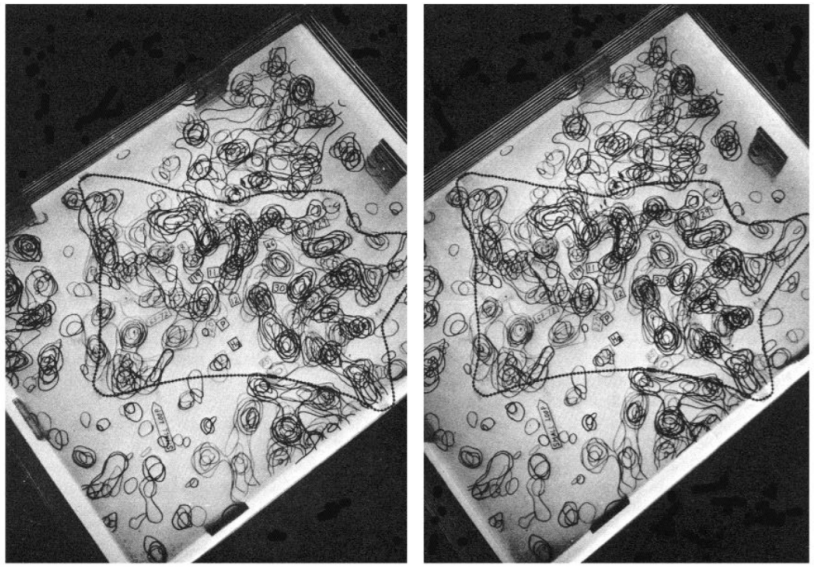
- **Academia**
- **Genex & DuPont CRD**
- **3DP**
- **Imiplex**

Early Protein Structures

The Structure of Ribonuclease-S at 3.5 A Resolution*

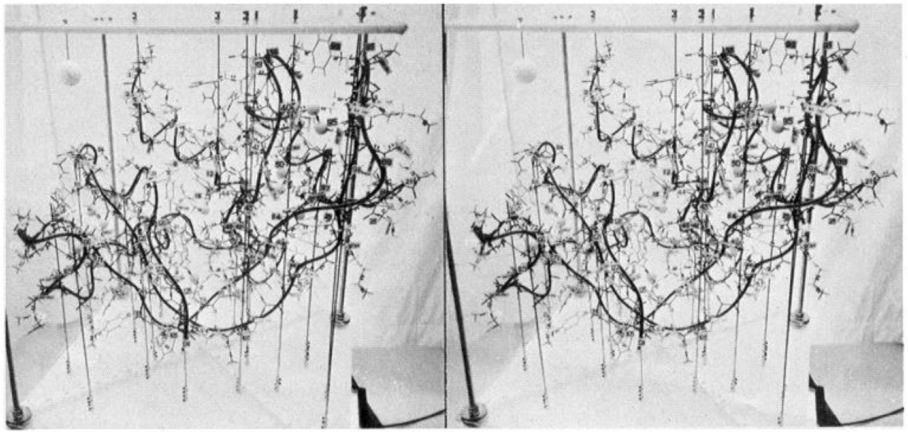
(Received for publication, May 31, 1967)

H. W. Wyckoff, Karl D. Hardman,‡ N. M. Allewell,
Tadashi Inagami,§ L. N. Johnson,¶
and Frederic M. Richards



Acknowledgments—We wish to extend special thanks to F. Raymond Salemme and J. David Weinland for technical assistance during this project.

WERMS in 2001

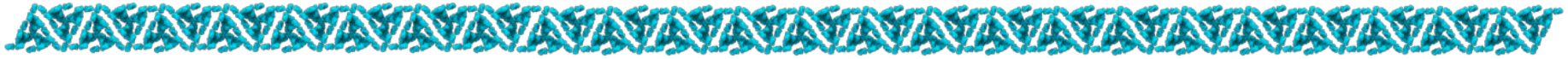


BA Molecular Biophysics
1962-67 Yale

IMIPLEX NanoSystems Inc

Early Structures

Protein Engineering: Globular Structure



- **3D structures of First Proteins (Myoglobin & Hemoglobin) showed organization of regular α-helices**
- **Lysozyme and Ribonuclease were more "complicated"**
- **β-sheets in globular proteins were complicated distorted structures, unlike regular, "flat" β-sheet structures proposed for silk and keratin**

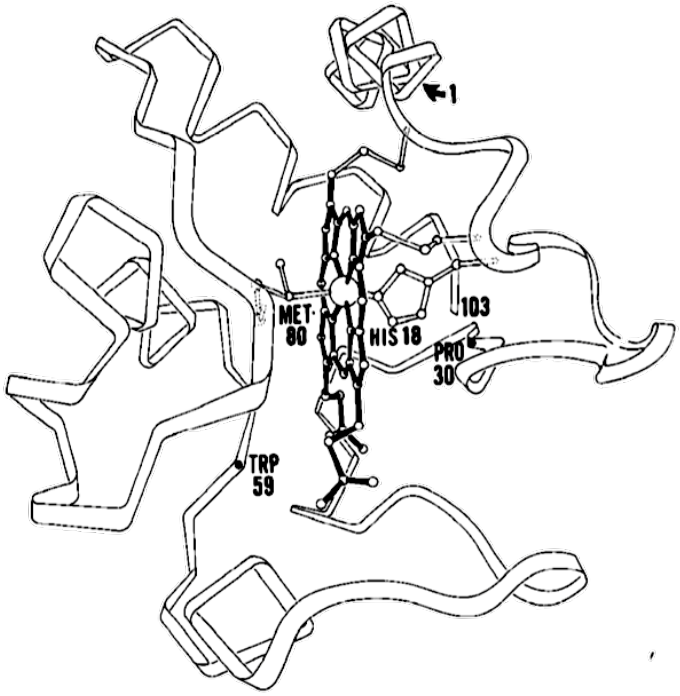
Structure & AA Sequence

STRUCTURE AND FUNCTION
OF CYTOCHROMES C

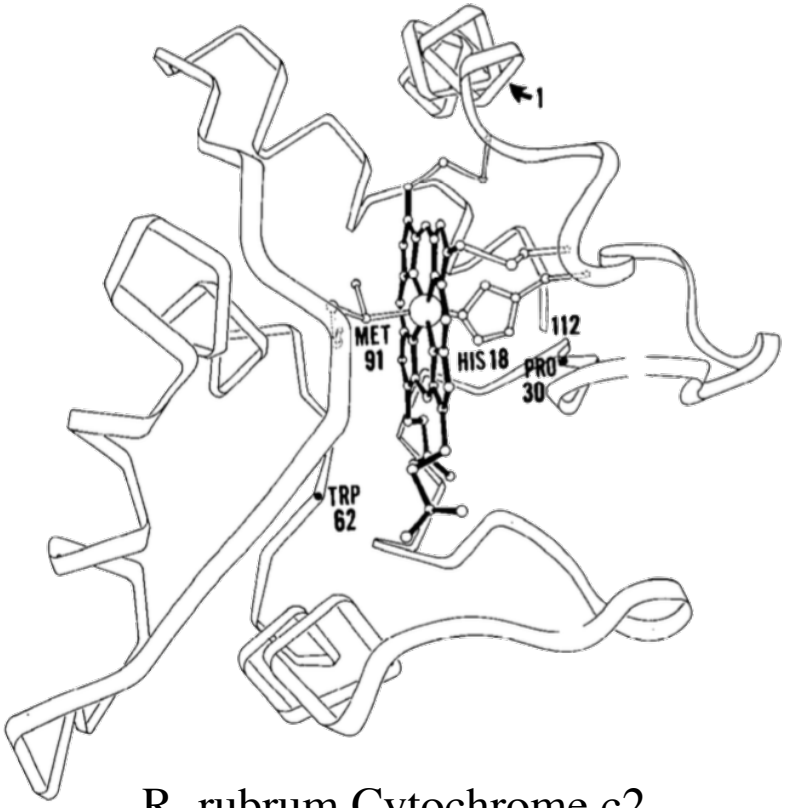
F. R. Salemme

Ann. Rev. Biochem. 1977. 46:299–329

MS, PhD Chemistry
UCSD 1968-72



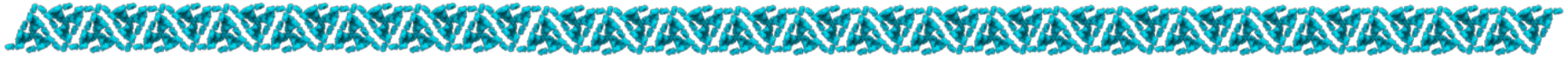
Tuna Cytochrome c



R. rubrum Cytochrome c2

IMIPLEX NanoSystems Inc

Protein Engineering: Structure & Homology



- **Mammalian and photosynthetic cytochromes functionally diverged ~ 2 Billion Years ago**
- **Amino Acid sequences show "twilight" similarity**

Q: How similar will the 3D structures be?

A: Overall 3D structure and key heme interactions are highly conserved

=> 3D structure more highly conserved than amino acid sequence

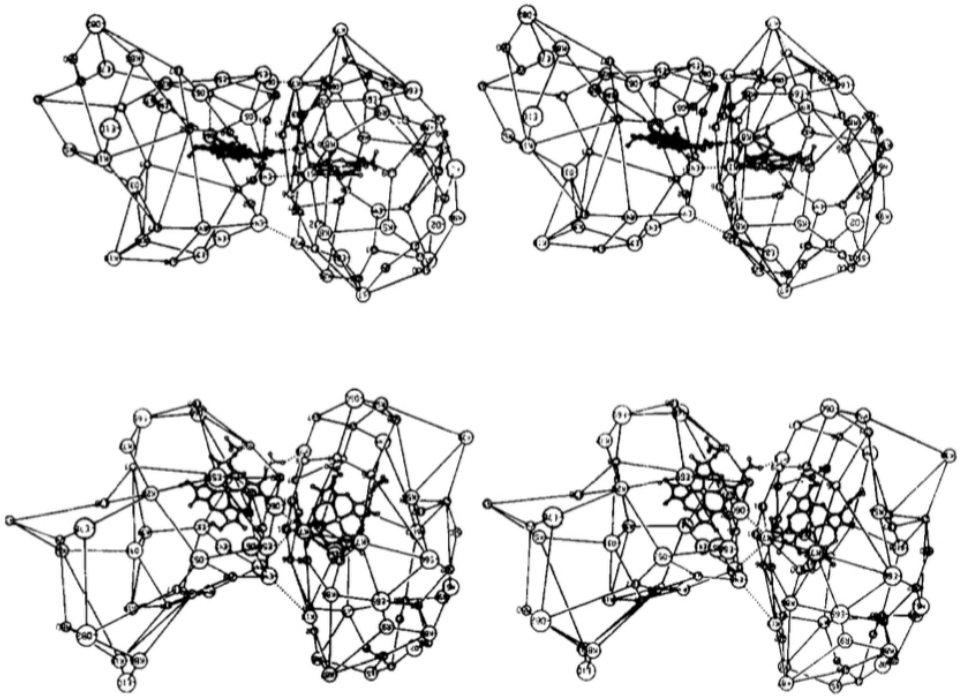
Intermolecular Recognition

An Hypothetical Structure for an Intermolecular Electron Transfer Complex of Cytochromes *c* and b_5

F. R. Salemme

J. Mol. Biol. (1976) **102**, 563–568

A CYTOCHROME c–b_5 INTERMOLECULAR COMPLEX 567



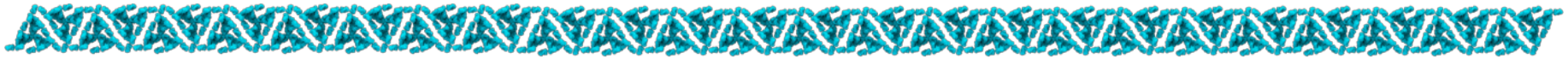
Department of Chemistry & Biochemistry
University of Arizona 1973-83

DuPont Central Research
1985-1991

1987

IMIPLEX NanoSystems Inc

Protein Engineering: Intermolecular Recognition



- **Cytochrome C and Cytochrome b5 known to efficiently undergo electron transfer reactions**
- **ET rate observed to be dependent on ionic strength**

=> Interaction mediated through complementary electrostatic interactions

Q: Can interheme ET take place through classical outer sphere mechanism (aka through direct heme contact)?

A: No, heme-heme closest approach is ~8Angstroms => ET by tunneling mechanism

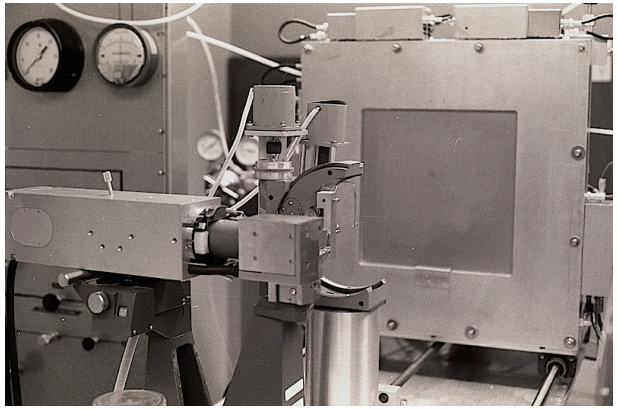
First published protein-protein model complex (pre interactive computer graphics)

Still working on the problem 10 years later

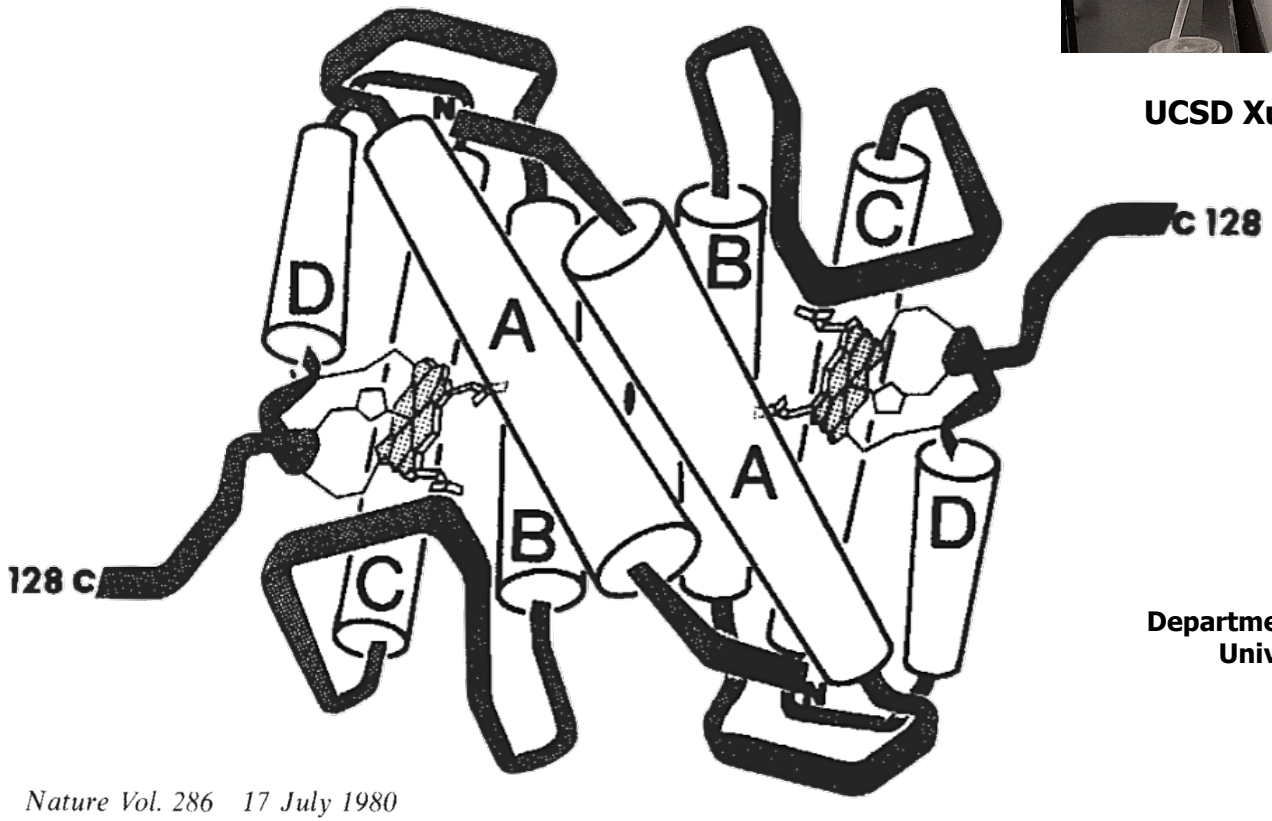
Recurrent Structural Motifs

Structure of cytochrome *c′*: a dimeric, high-spin haem protein

Patricia C. Weber*§, R. G. Bartsch†, M. A. Cusanovich*, R. C. Hamlin‡, A. Howard‡, S. R. Jordan*, M. D. Kamen†, T. E. Meyer†, D. W. Weatherford*, Nguyen huu Xuong‡ & F. R. Salemme*



UCSD Xuong 2D X-Ray Detector



Nature Vol. 286 17 July 1980

Department of Chemistry & Biochemistry
University of Arizona 1973-83

IMIPLEX NanoSystems Inc

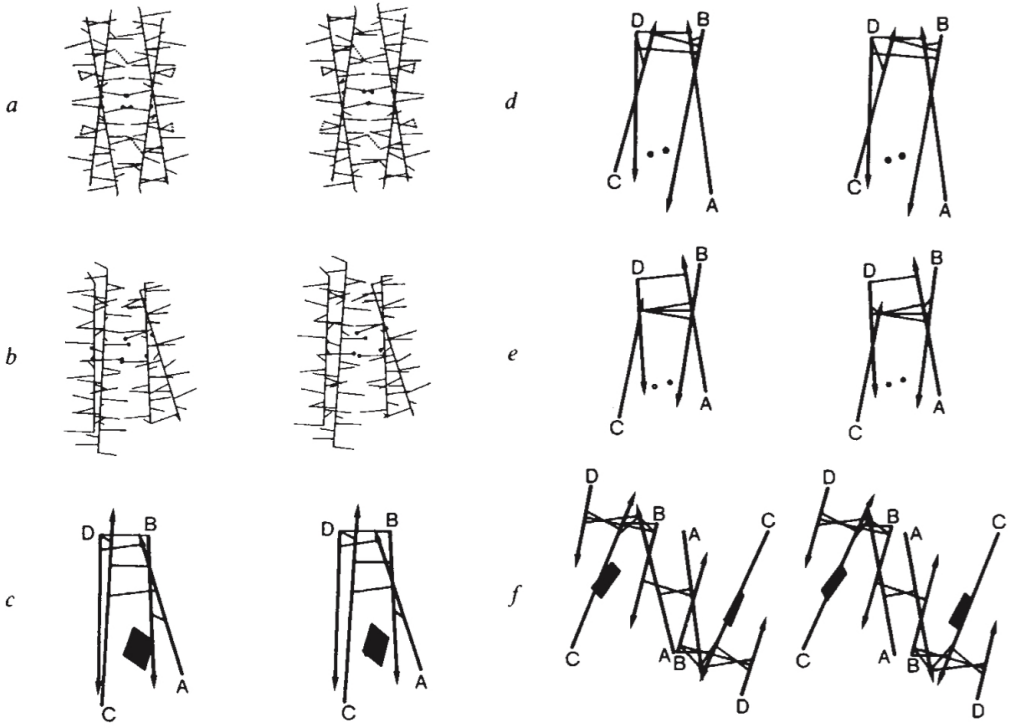
Recurrent Structural Motifs

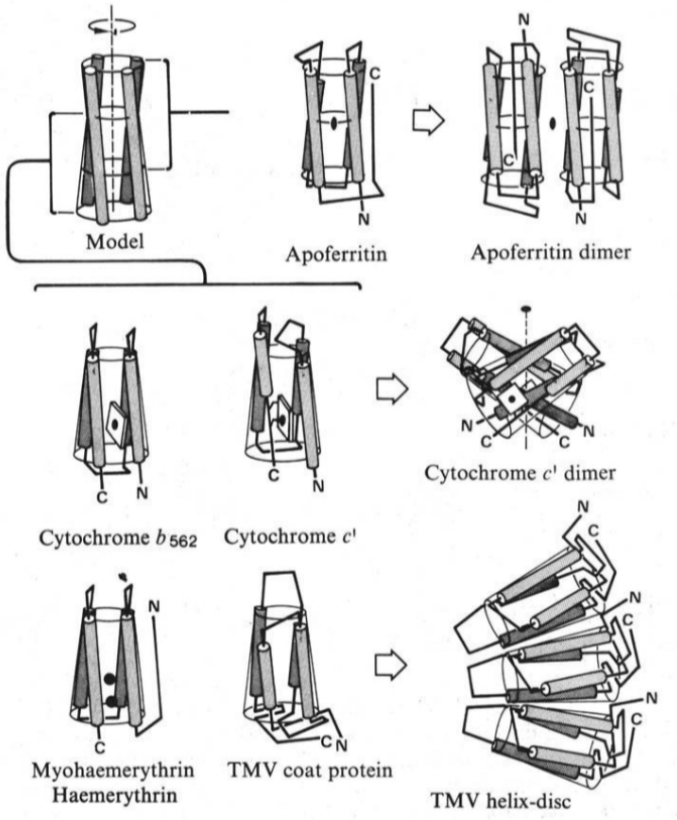
Structural and functional diversity in 4-α-helical proteins

Patricia C. Weber & F. R. Salemme

Department of Biochemistry, New Chemistry Building, University of Arizona, Tucson, Arizona 85721

Department of Chemistry & Biochemistry
University of Arizona 1973-83





Nature Vol. 287 4 September 1980

IMIPLEX NanoSystems Inc

Proc. Natl. Acad. Sci. USA
Vol. 79, pp. 4545–4549, August 1982
Biochemistry

Recurrent Structural Motifs

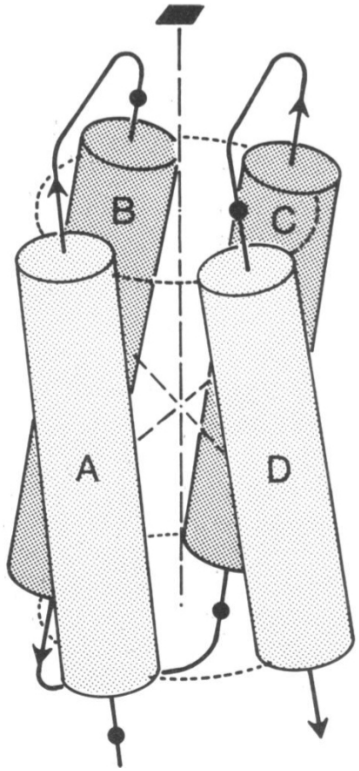
α-Helix dipole model and electrostatic stabilization of 4-α-helical proteins

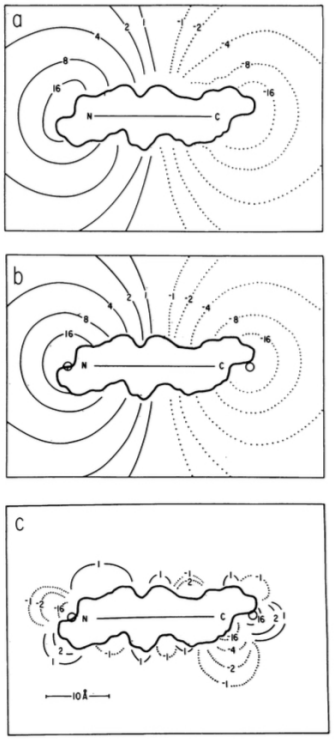
(electrostatic interactions/protein structure)

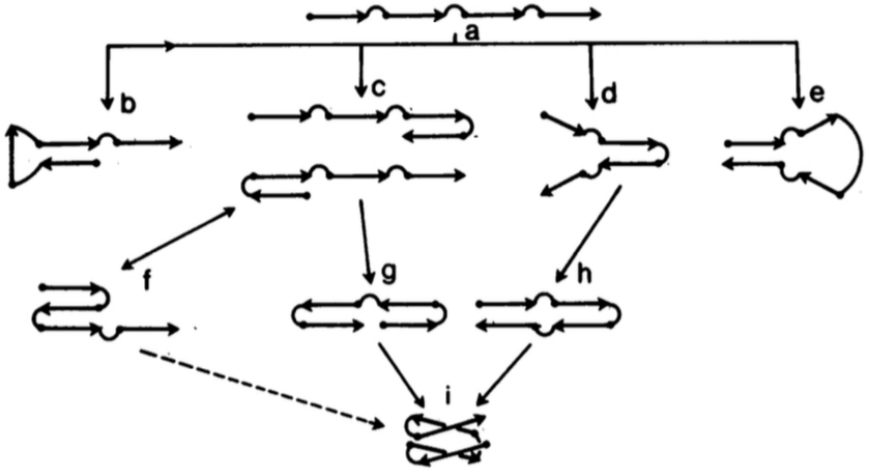
ROBERT P. SHERIDAN*, RONALD M. LEVY*†, AND F. R. SALEMME‡§

*Department of Chemistry, Rutgers University, New Brunswick, New Jersey 08903; and ‡Department of Biochemistry, University of Tucson, Tucson, Arizona 85717

Communicated by Frederic M. Richards, April 12, 1982

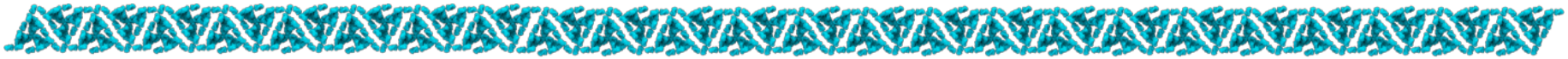






IMIPLEX NanoSystems Inc

Protein Engineering: 4-α-Helical Bundle Proteins



- **Recurrent Structural Motif seen among several proteins with no apparent sequence homology (TMV, hemerythrin, myohemerhthrin, Cyt c′, Cyt b562)**

Q: Related by convergent or divergent evolution?

A: All of the structures are organized on the same principles, which results from a degeneracy in possible side chain packing arrangements for 4 α-helices crossing at ~18 degs.

⇒ **Convergent evolution**

⇒ **Helix macrodipoles can facilitate folding and stabilize final structure**

⇒ **Structural motifs recur because they represent stable local minima on protein energy landscapes**

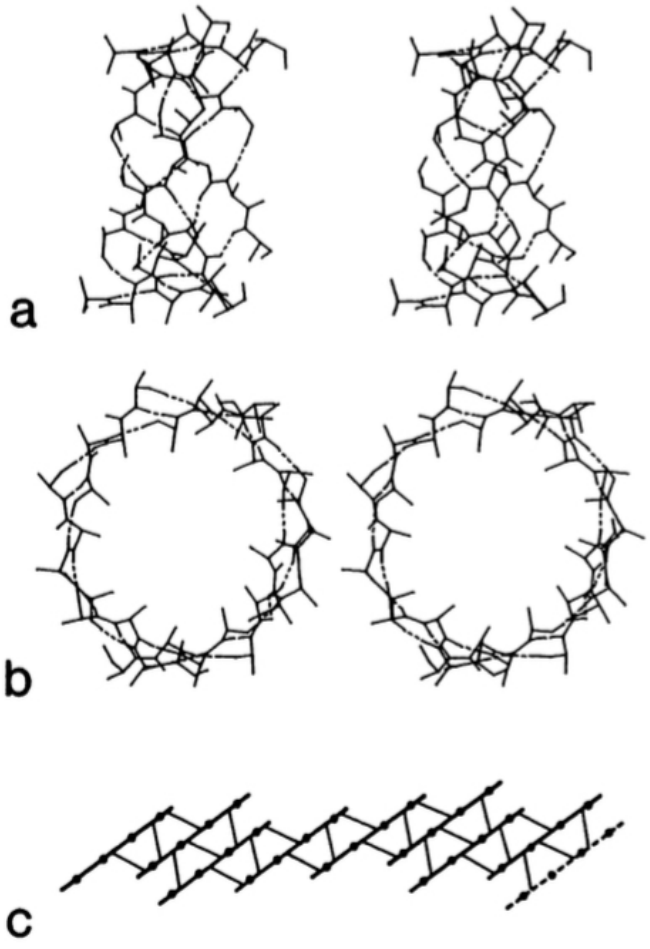
Recurrent Structural Motifs

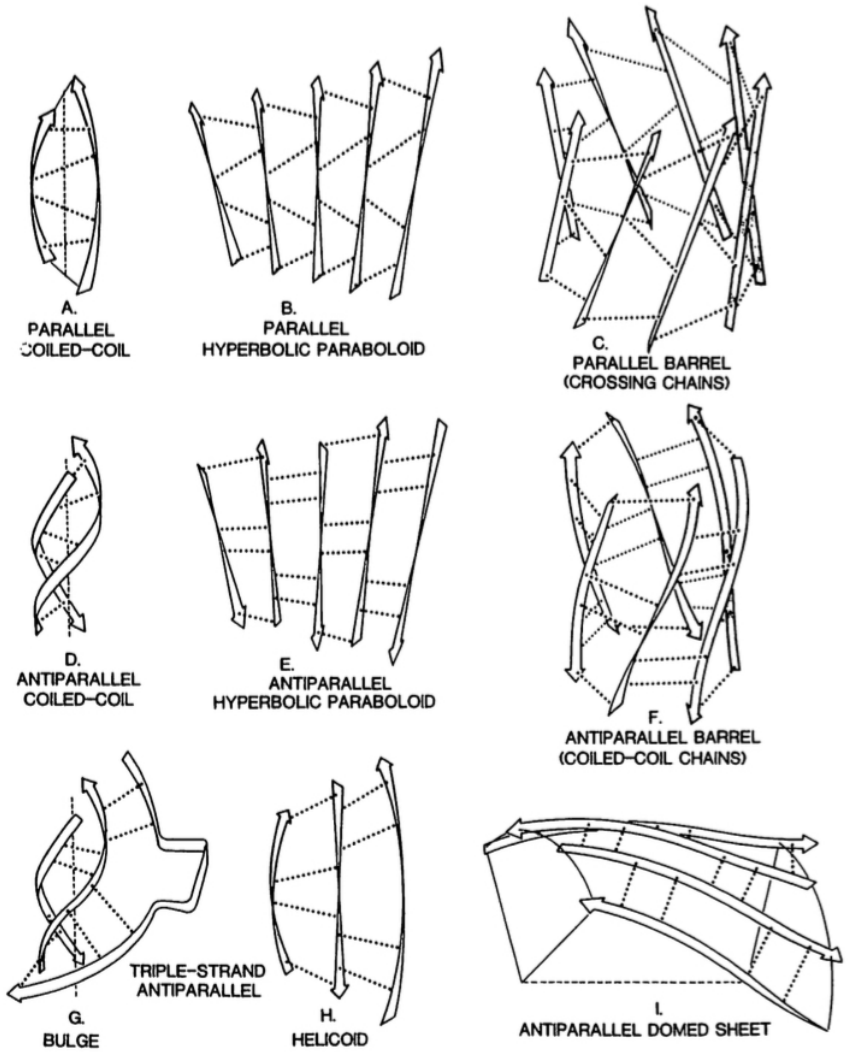
STRUCTURAL PROPERTIES OF PROTEIN β-SHEETS

F. R. SALEMME

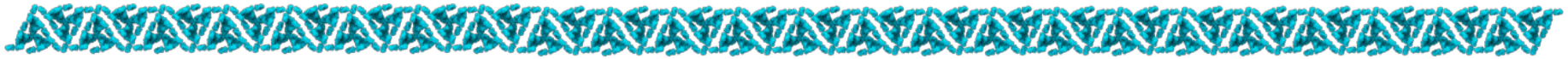
Department of Chemistry & Biochemistry
University of Arizona 1973-83

Prog. Biophys. molec. Biol., Vol. 42, pp. 95-133, 1983.





Protein Engineering: β-Sheet Geometry



- **β-sheets in proteins have unusual looking curved shapes**

Q: Where do these complex shapes come from?

A: Basic properties can all be explained with atomic precision as the results of surface area minimization under the constraints of interchain H-Bond geometry

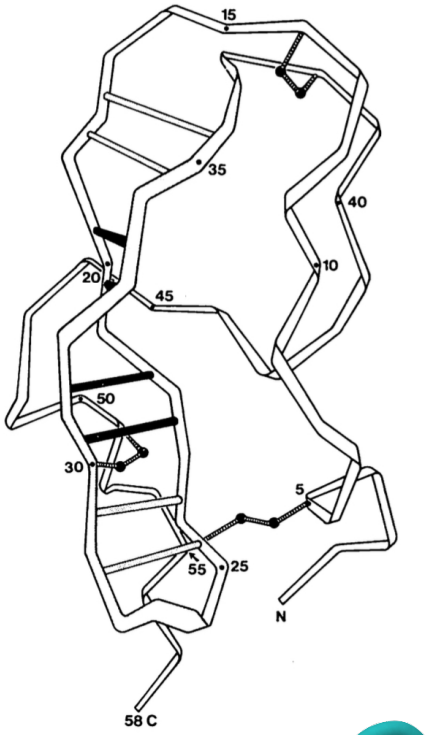
=> Structural motifs recur because they represent stable local minima on protein energy landscapes

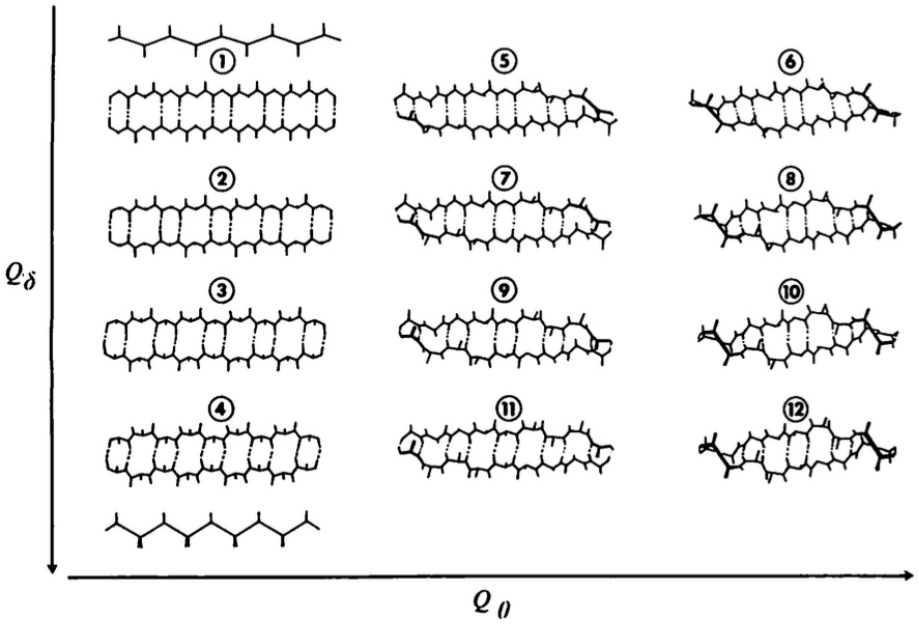
Yale University Sabbatical
1981-82

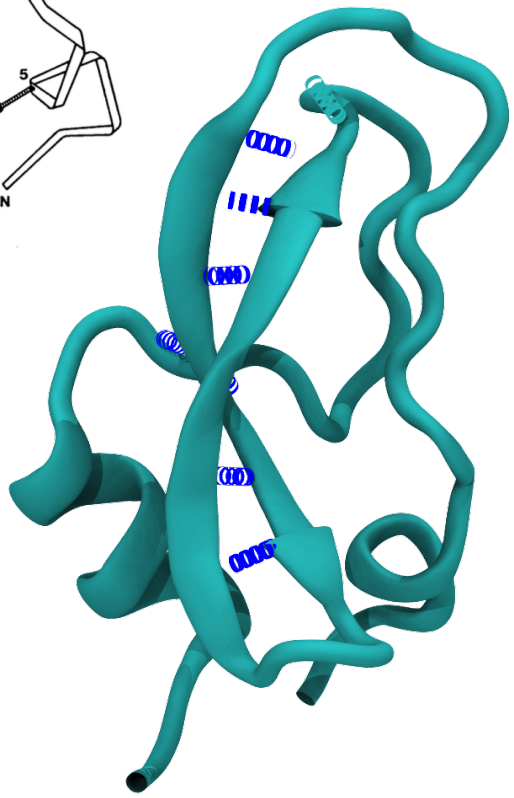
Cooperative motion and hydrogen exchange stability in protein β-sheets

F. R. Salemme*

Department of Molecular Biophysics and Biochemistry,
Yale University, New Haven, Connecticut 06511, USA







Reprinted from Nature, Vol. 299, No. 5885, pp. 754-756, 21 October 1982

IMIPLEX NanoSystems Inc

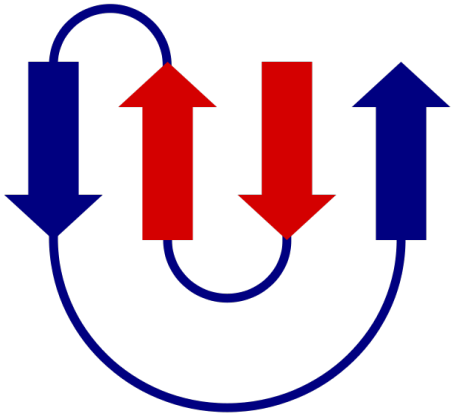
Protein Engineering: β-Sheet Geometry & Folding

- **Some structures in proteins show unusual amide proton exchange stability, despite high level of solvent accessibility**

Q: How is this possible?

A: Some structures, like double strand antiparallel β-sheets are intrinsically cooperative in their flexibility.

=> Accounts for recurrent patterns of Greek Key folding motifs in proteins



Protein Engineering: A Personal Perspective

- **Academia**
- **Genex & DuPont CRD**
- **3DP**
- **Imiplex**

Protein Engineering

Kevin. M. Ulmer (Genex Corporation)

Science March 1983

Summary. The prospects for protein engineering, including the roles of x-ray crystallography, chemical synthesis of DNA, and computer modeling of protein structure and folding, are discussed. It is now possible to attempt to modify many different properties of proteins by combining information on crystal structure and protein chemistry with artificial gene synthesis. Such techniques offer the potential for altering protein structure and function in ways not possible by any other method.

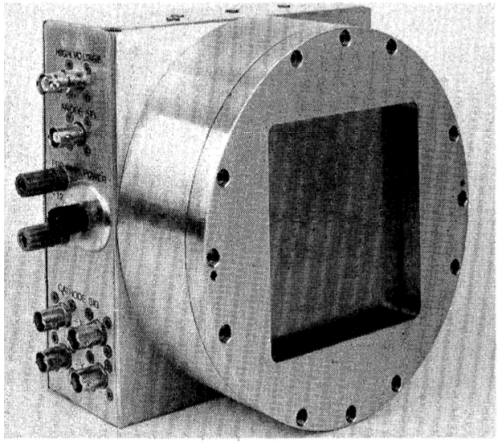


Fig. 1. Electronic position-sensitive x-ray detector. [Courtesy of Xentronics Company, Inc., Cambridge, Massachusetts]

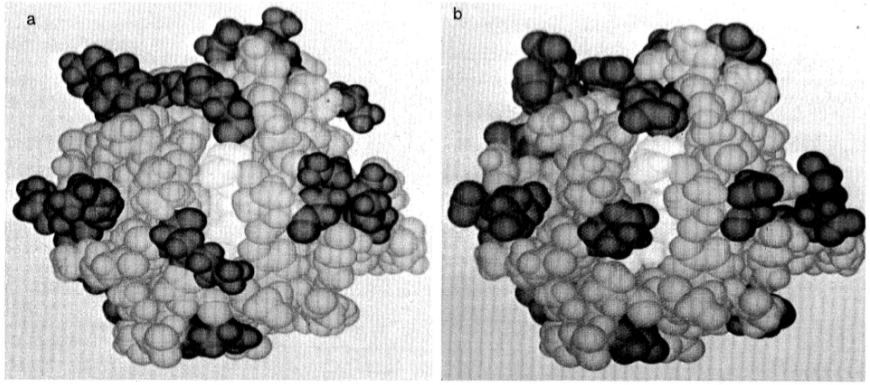


Fig. 2. Application of interactive three-dimensional computer graphics with a molecular model of tuna cytochrome c. (a) Native structure with positively charged lysine residues indicated by dark shading. (b) Lysine residues have been graphically replaced with negatively charged glutamic acid residues to simulate a protein engineering experiment that might reverse the surface charge of the protein. [Courtesy of R. J. Feldmann, National Institutes of Health, Bethesda, Maryland]

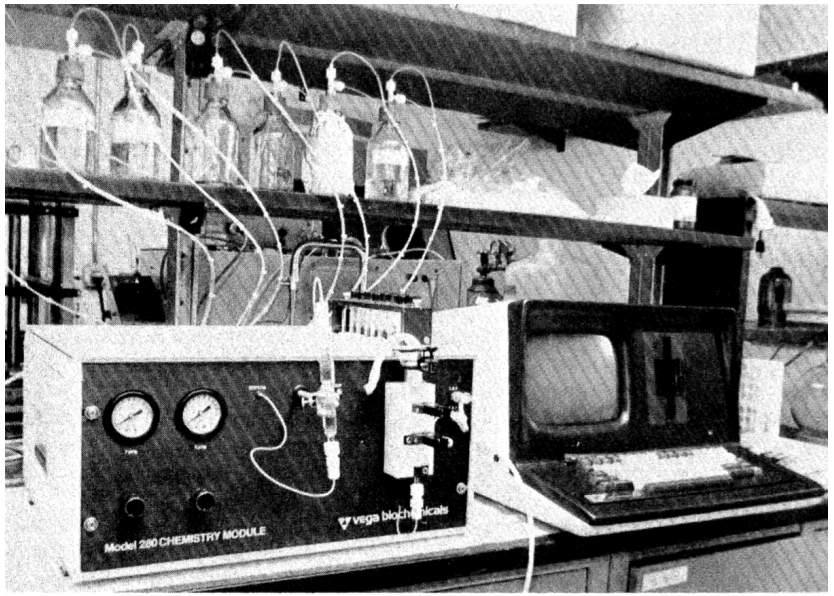


Fig. 3. Automated instrumentation for the synthesis of oligonucleotides.

Genex Corporation
1983-84

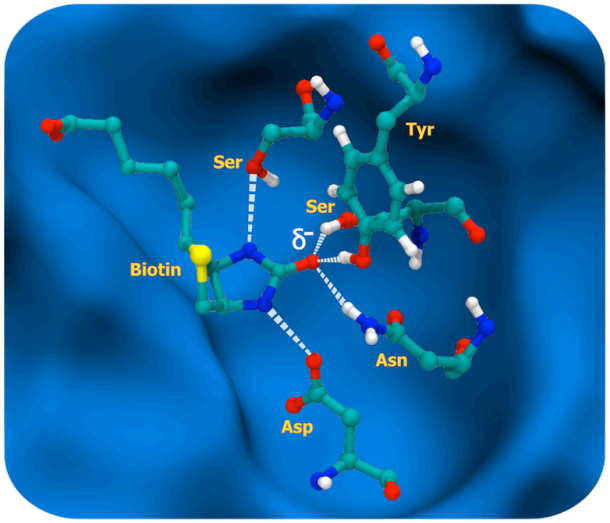
IMIPLEX NanoSystems Inc

Streptavidin

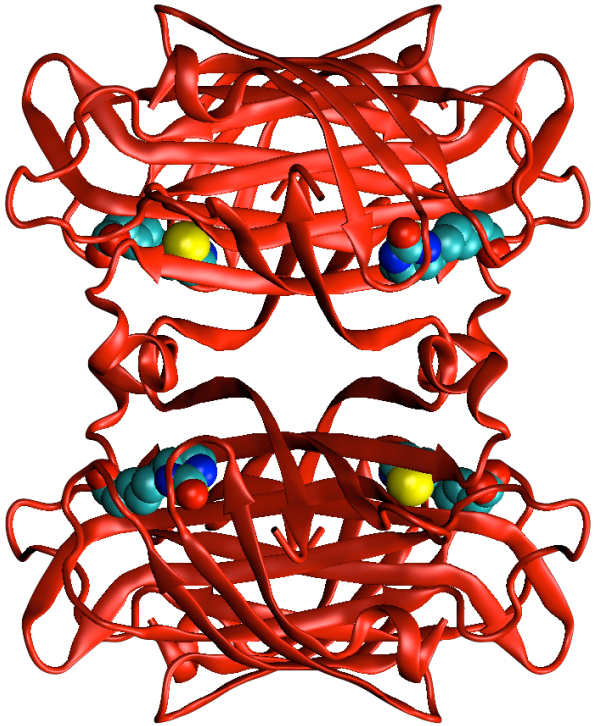
Structural Origins of High-Affinity Biotin Binding to Streptavidin

PATRICIA C. WEBER, D. H. OHLENDORF, J. J. WENDOLOSKI, F. R. SALEMME*

The high affinity of the noncovalent interaction between biotin and streptavidin forms the basis for many diagnostic assays that require the formation of an irreversible and specific linkage between biological macromolecules. Comparison of the refined crystal structures of apo and a streptavidin:biotin complex shows that the high affinity results from several factors. These factors include the formation of multiple hydrogen bonds and van der Waals interactions between biotin and the protein, together with the ordering of surface polypeptide loops that bury the biotin in the protein interior. Structural alterations at the biotin binding site produce quaternary changes in the streptavidin tetramer. These changes apparently propagate through cooperative deformations in the twisted β sheets that link tetramer subunits.



DuPont Central Research
1985-1991

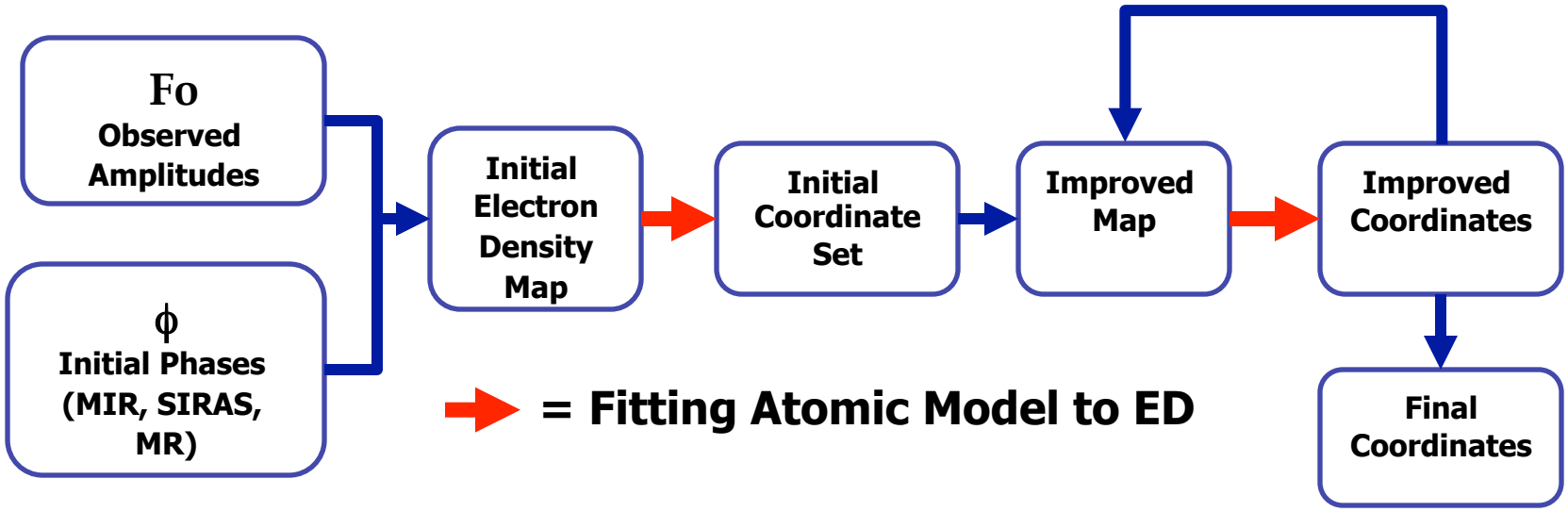


1244 Literature Citations as of 03/18

Structural Origins of High Affinity Biotin Binding to Streptavidin, P.C. Weber, D.H. Ohlendorf, J.J. Wendoloski, F.R. Salemme, ***Science*** 1989; 243: 85-88

IMIPLEX NANOSYSTEMS INC

Protein Refinement
via Protein Structural Heuristics



What you know:

- **Electron density is positive or zero (never negative)**
- **Peptide geometry is regular**
- **Secondary structure geometry is regular (for the most part)**
- ***Proteins are assembled from recurring "foldon" structural motifs – often having an identifiable amino acid sequence signature***

Molecular Modeling With Substructure Libraries Derived From Known Protein Structures

Barry C. Finzel, S. Kimatian, D. H. Ohlendorf, J.J. Wendoloski, M. Levitt*, and F.R. Salemme

DuPont Central Research
1985-1991

FRAGL

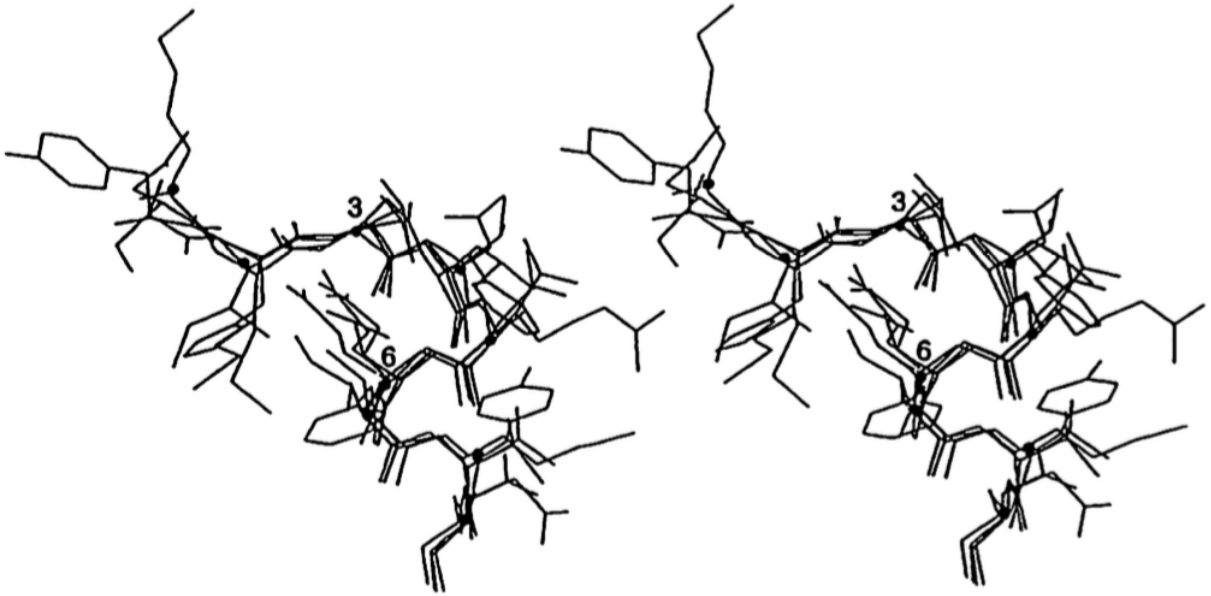
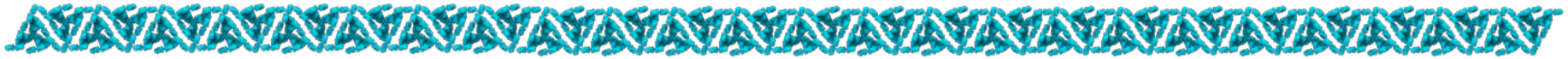


Figure 4. Stereoscopic representation of four similar protein fragments extracted from the library of known structures. The four (and the amino acid sequence) are from 1) Cytochrome P450 residues 190-198 (SMTFAEAKE); 2) Cytochrome c Peroxidase residues 162-170 (NMNDREVVA); 3) Carboxypeptidase residues 12-20 (YHTLDEIYD); and 4) Parvalbumin residues 96-104 (KIGVDEFTA). All are superimposed on the C-α backbone of Calmodulin residues 99-107 (FISAAELRH) (shown as dots) used as a target conformational template.

In Crystallographic and Modeling Methods in Molecular Design (S Ealick & C Bugg eds.) Springer Verlag, New York, 175-189 (1990)

IMIPLEX NanoSystems Inc

Protein Engineering: Foldons



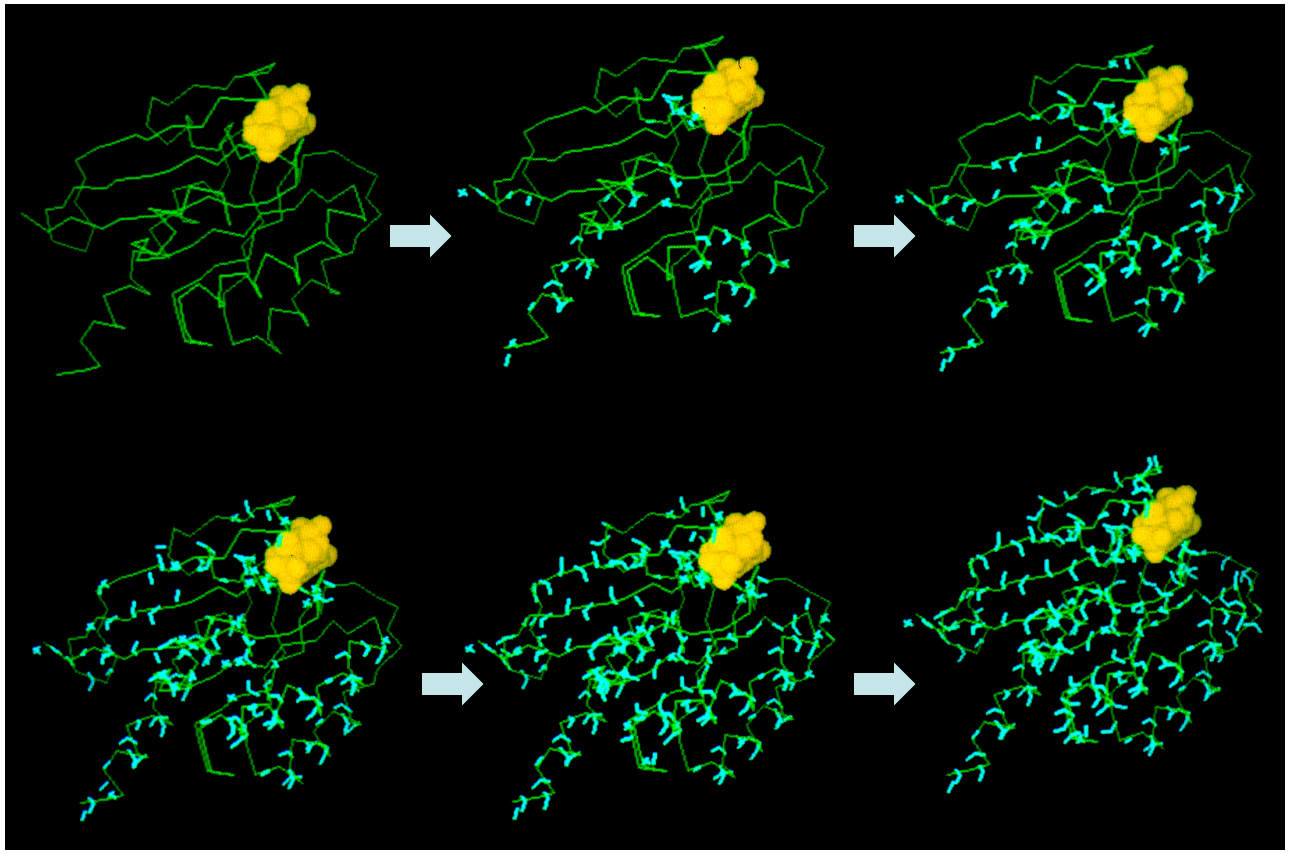
- **Local Recurrent Structural Motifs, despite only vestigial sequence similarity**
- **Often incorporate elements of secondary structure**
- **Implemented as FRAGL program (1990) for rapid construction of protein models for X-ray crystallography**
- **Ultimately the basis for heuristic "protein folding" algorithms (aka Rosetta)**

DuPont Central Research
1985-1991

PROBIT: A statistical approach to modeling proteins from partial coordinate data using substructure libraries

J.J. Wendoloski and F.R. Salemme

PROBIT

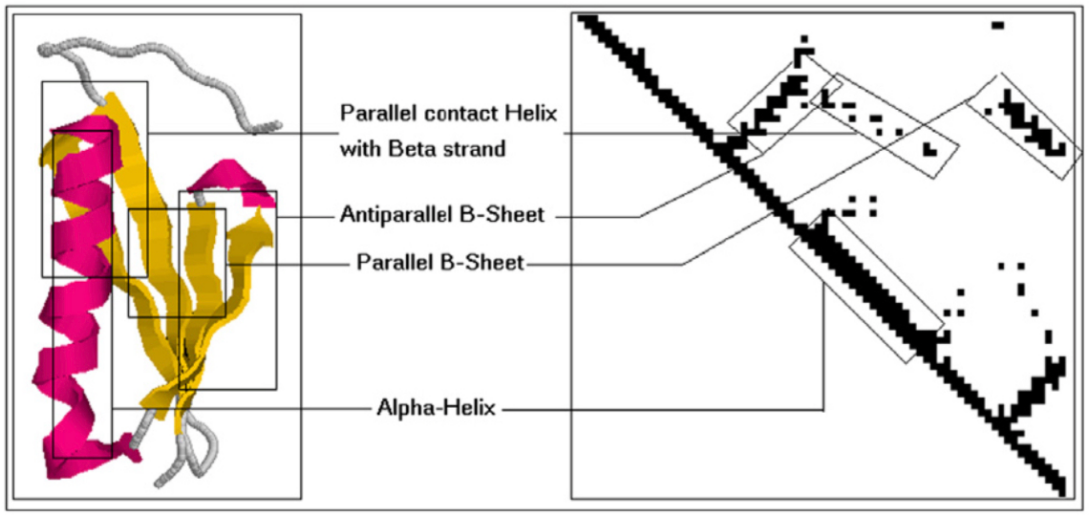


J. Mol. Graphics, 1992, Vol. 10, June

IMIPLEX NanoSystems Inc

Protein Engineering: Amino Acid Side Chain Rotamers

- **Many side-chain rotamers are fixed in a local secondary structure context**
- **Context-dependent, Probability-weighted side chain reconstruction implemented in PROBIT (1992) for rapid construction of protein models for X-ray crystallography**
- **Ultimately an important feature of heuristic "protein folding" and structural optimization algorithms (aka Rosetta)**



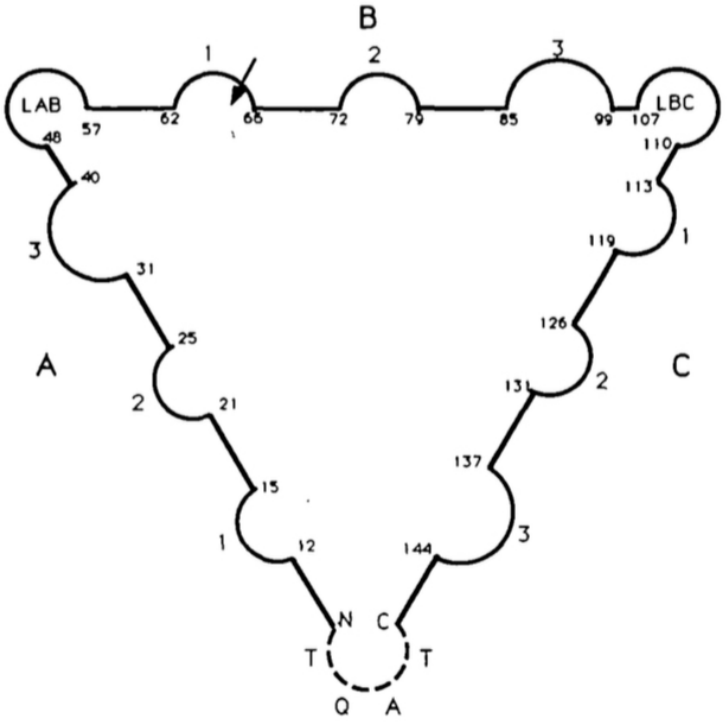
A.A. Abu-Doleh et al./Journal of Biomedical Informatics 45 (2012) 173–183

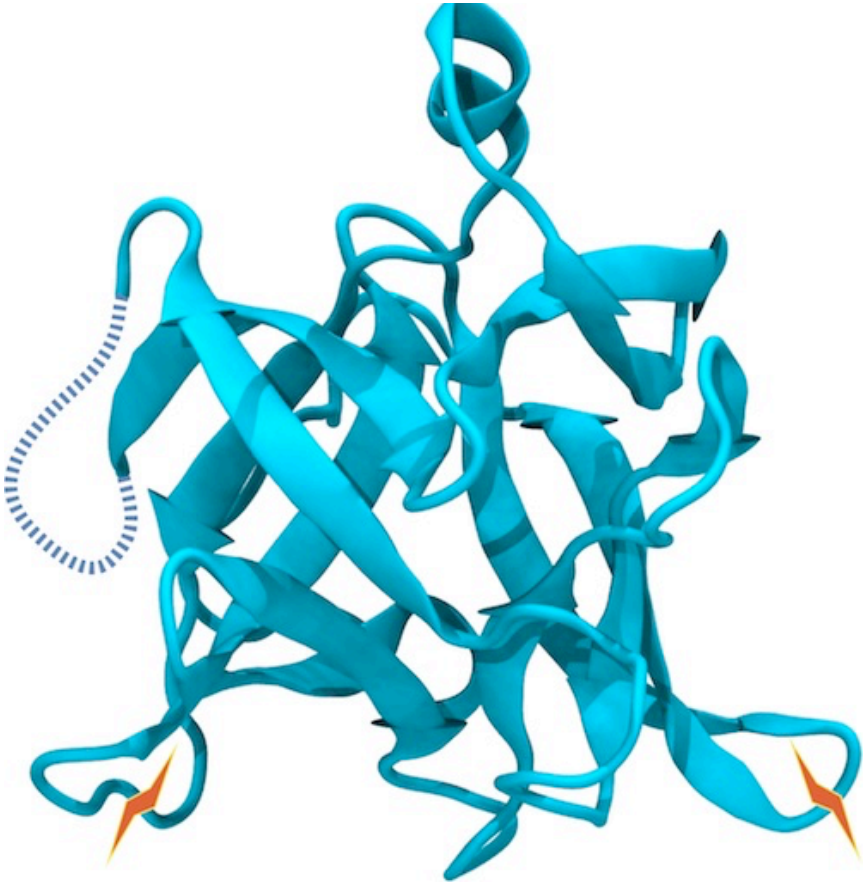
IMIPLEX NanoSystems Inc

Protein Engineering vol.5 no 5 pp.427–431, 1992

DuPont Central Research
1985-1991

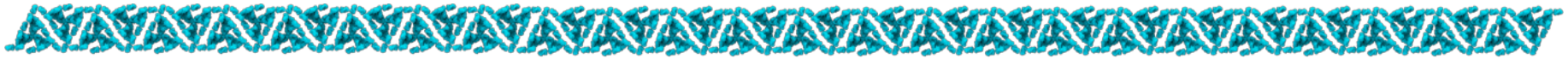
Permuteins of interleukin 1β—a simplified approach for the construction of permutated proteins having new termini





IMIPLEX NanoSystems Inc

Protein Engineering: Cyclic Permuteins

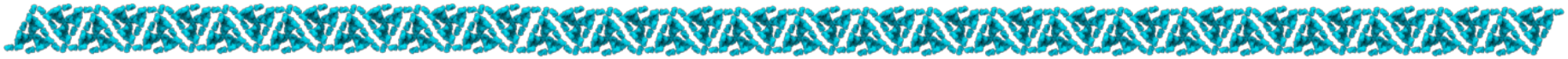


Q: Do proteins fold from one end of the polypeptide chain?

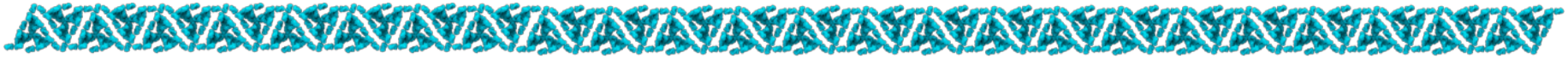
A: For IL1 (a 3-domain β-trefoil motif), N and C terminus could be located at any of three domain boundaries,. With no apparent adverse effect on folding or function

=> Reconnection or grafting of protein domains should not impede proper folding for suitable stable folding protein structural motifs.

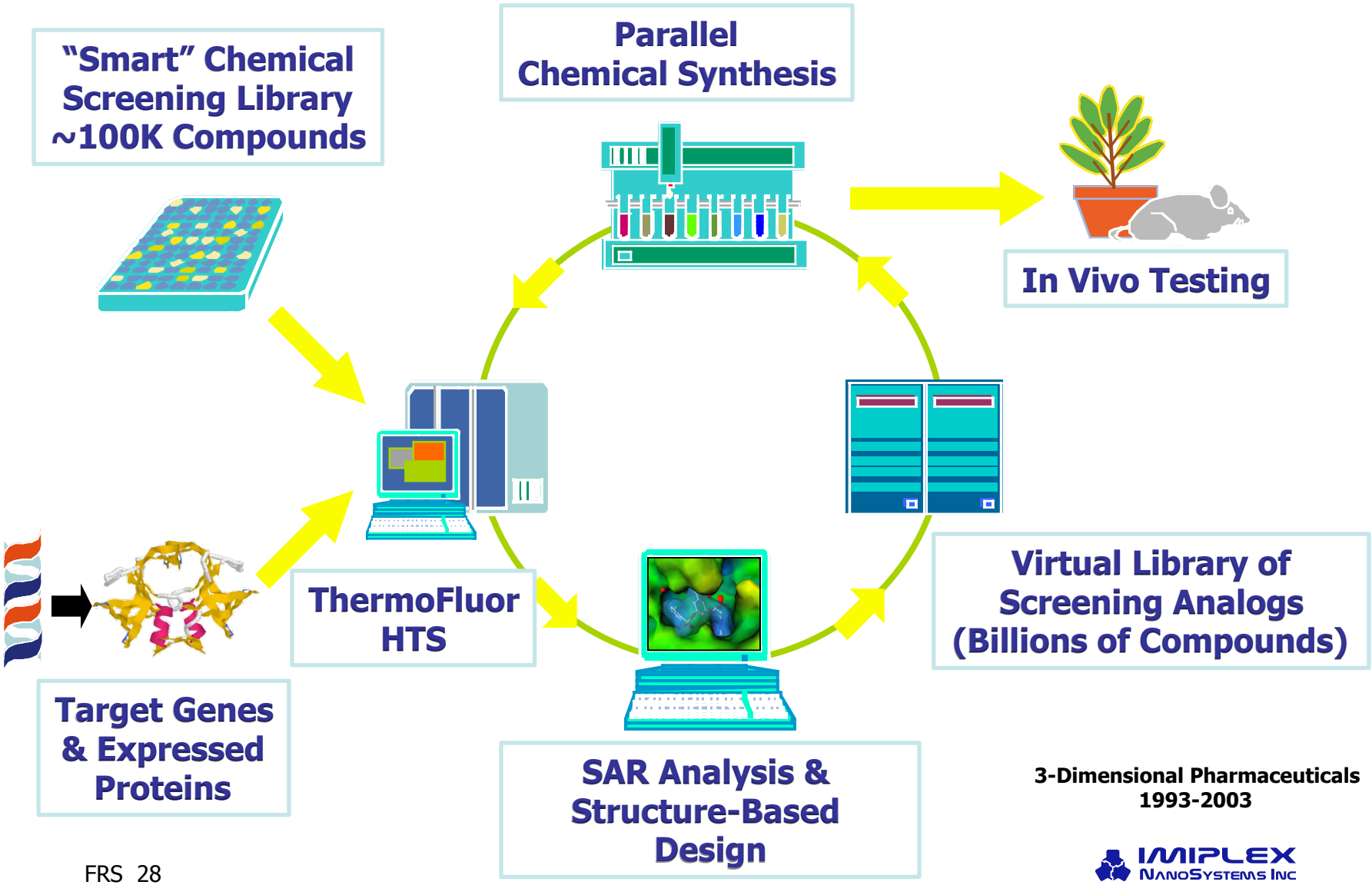
Protein Engineering: A Personal Perspective



- **Academia**
- **Genex & DuPont CRD**
- **3DP**
- **Imiplex**



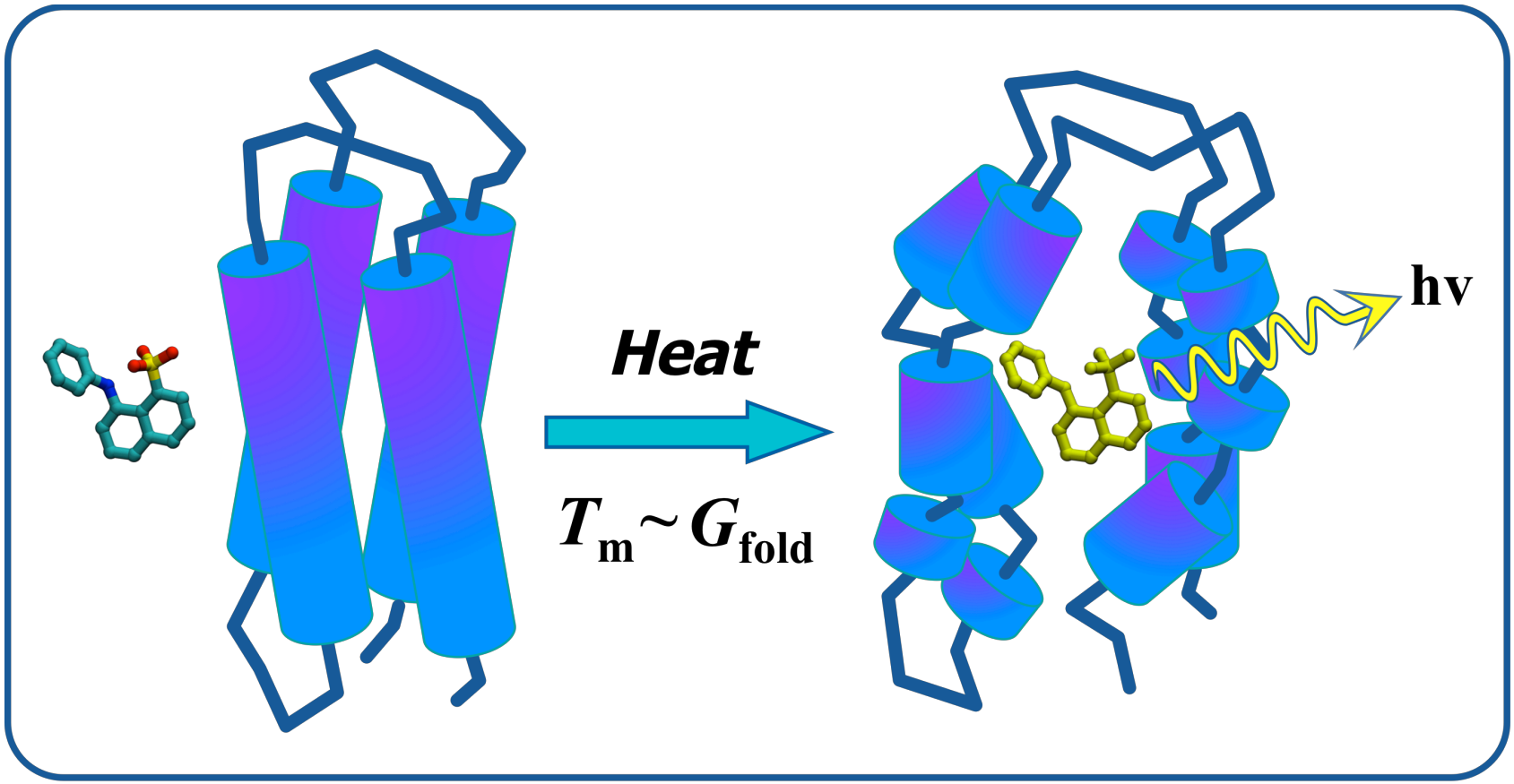
3DP Drug Discovery Platform



3-Dimensional Pharmaceuticals
1993-2003

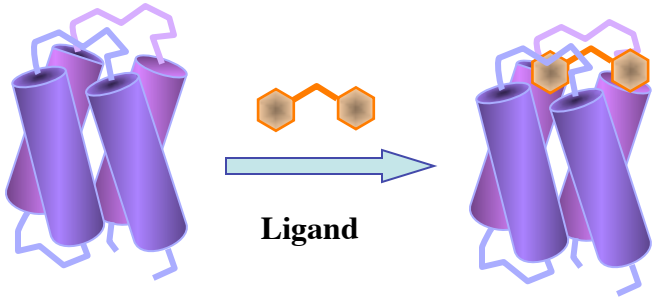
IMIPLEX NanoSystems Inc

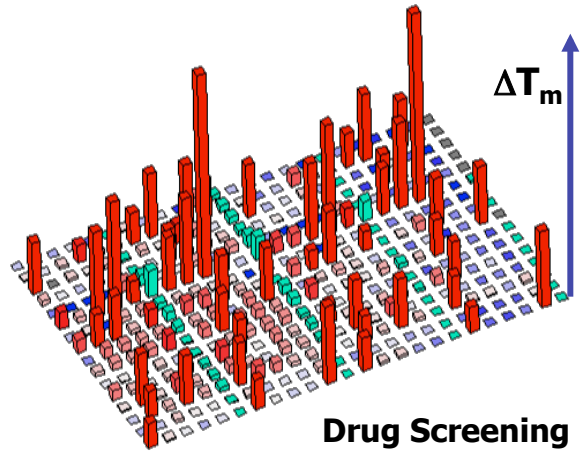
Thermofluor (aka DSF)

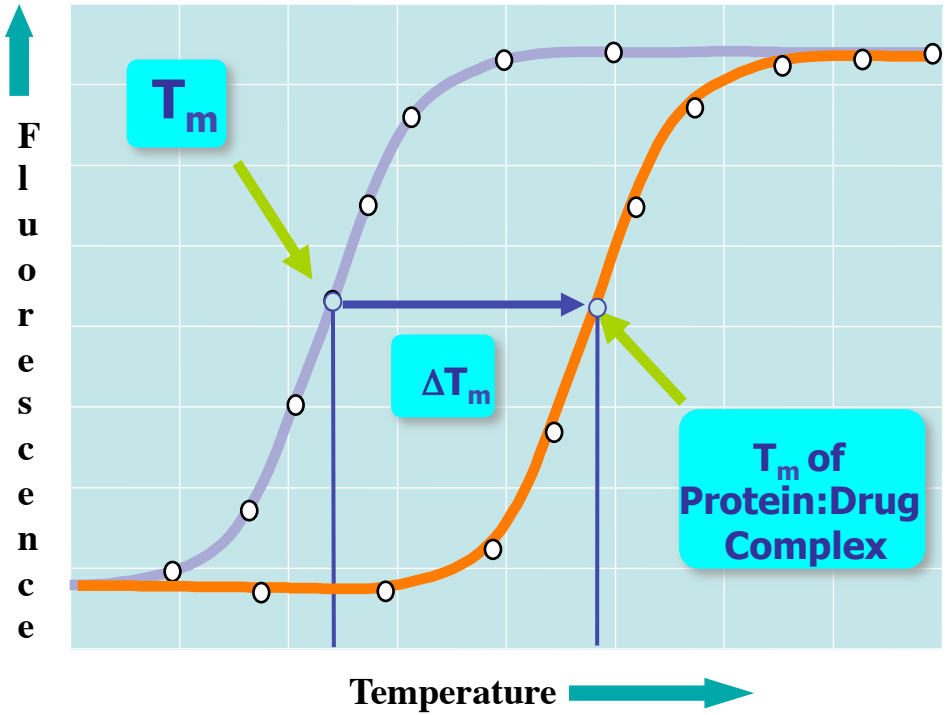


- **Thermofluor gives an optical readout of protein melting as e.g. observed with DSC.**
- **Optical readout is much more sensitive than direct thermal measurement.**
- **Allows parallel measurements in 384 well-plates using fluorescent imaging plate readers.**

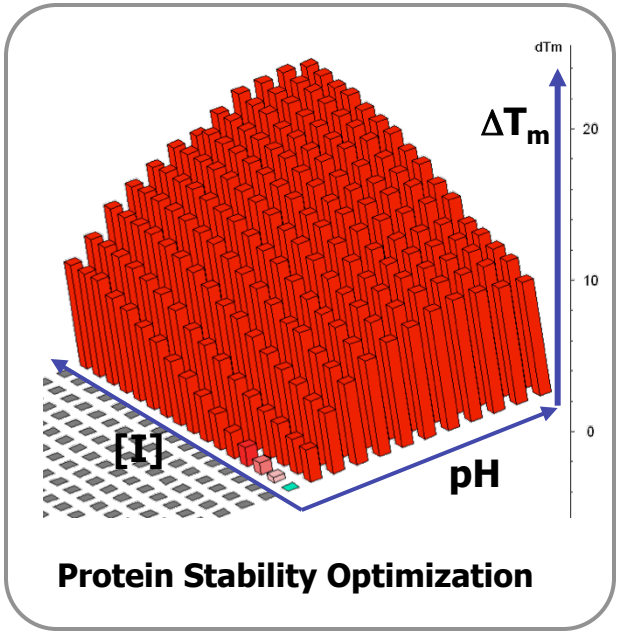
Ligand Binding Can Stabilize Protein Structure







But Also …

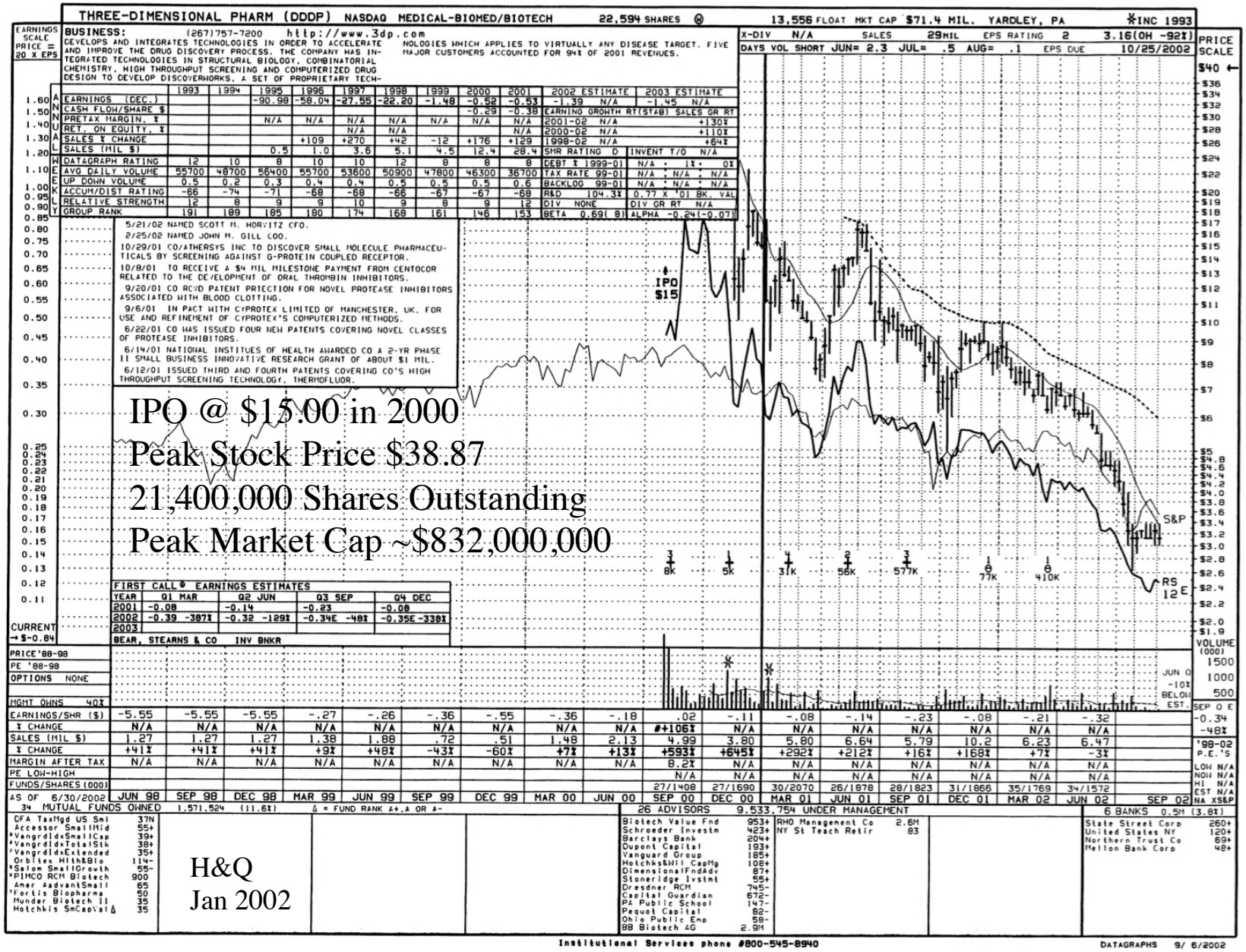


3-Dimensiaoal Pharmaceuticals
1993-2003

IMIPLEX NanoSystems Inc

3-Dimensiaoal Pharmaceuticals
1993-2003

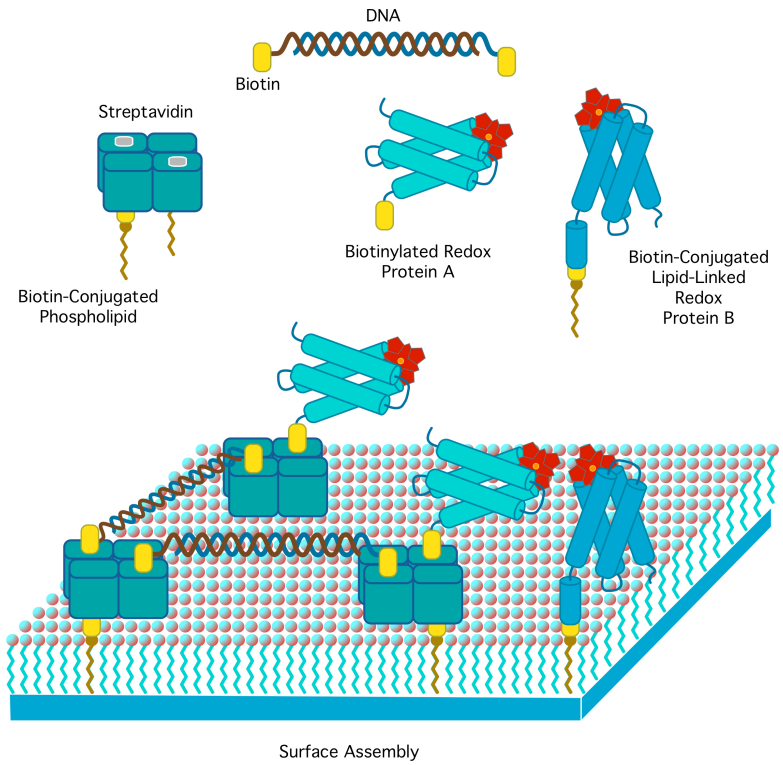
3DP (*Nasdaq* DDDP) Stock Price History



Protein Engineering: A Personal Perspective

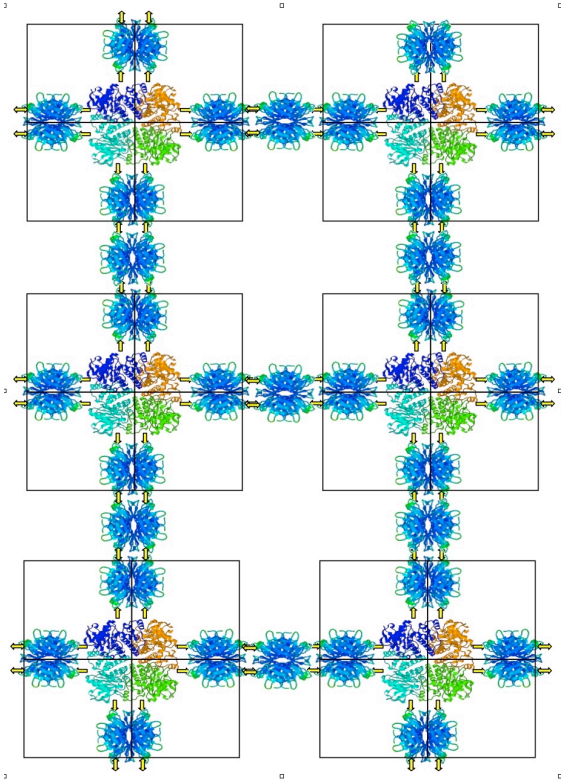
- **Academia**
- **Genex & DuPont CRD**
- **3DP**
- **Imiplex**
  - **Protein Engineering for Metamaterials**

Precedents



Streptavidin-Linked Nanostructures for Molecular Electronics
Salemme & Sligar, 1992

DuPont Central Research
1985-1991



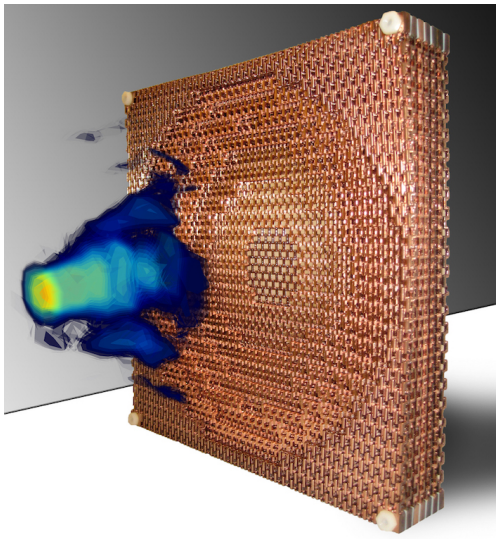
Streptavidin-Linked 2D Protein Lattices
Ringler & Schulz, 2003

Imiplex LLC 2003

IMIPLEX NanoSystems Inc

Metamaterials & Protein-Based Nanostructure

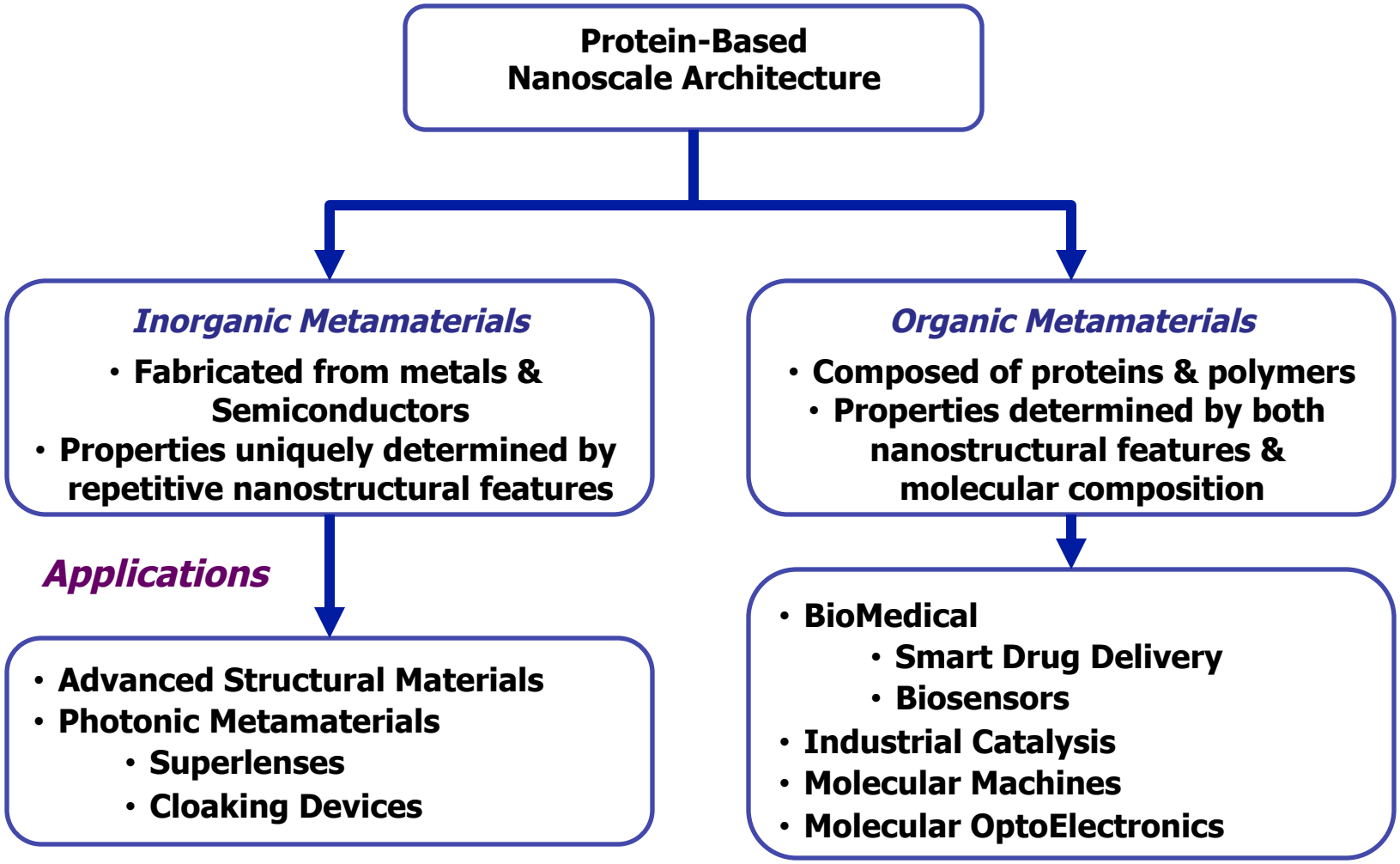
- **Metamaterials are new forms of matter - not found in nature - that derive their novel properties from the precise nanoscale structure and organization of their components.**
- ***Inorganic Metamaterials,* composed of metal or ceramic, have demonstrated ability to modulate electromagnetic radiation in ways that are not possible using conventional materials. Availability of scalable manufacturing methods remain key limitation to commercialization.**
- ***Organic Metamaterials,* composed of engineered proteins, represent a new class of materials with diverse applications for medicine, biosensors, molecular electronics, energy, and industrial-scale chemical processes.**



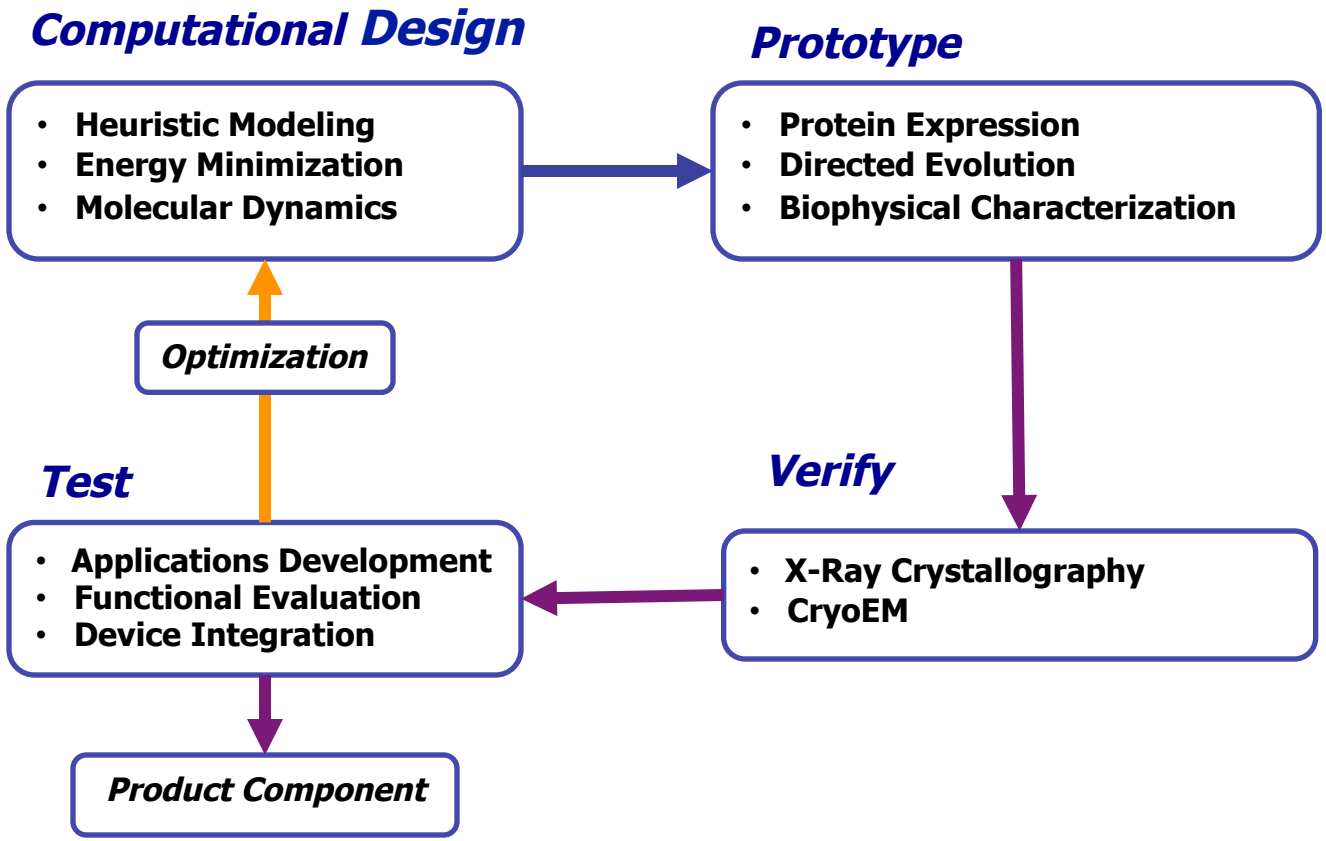
Protein-Based Nanotechnology

- Spontaneous self-assembly with atomic precision
- Extensive range of natural functionality
  - Catalysis
  - Chemosensation
  - Charge Separation
  - Photosynthesis
  - Mechanical Motion
- ~138,000 3D structures (~75,000 unique) known from protein crystallography (many more sequences)
- Many highly stable structures (>90 deg C)
- Protein nanostructural components are readily engineered using computational tools and recombinant DNA technology
- Natural interface to biological systems
- Allows hybrid bottom-up & top-down self-assembly processes
- Feasible routes to large-scale, low-cost production
- Provides basis for non-organic meta-material formation

Protein-Based Nanotechnology for Metamaterial Applications



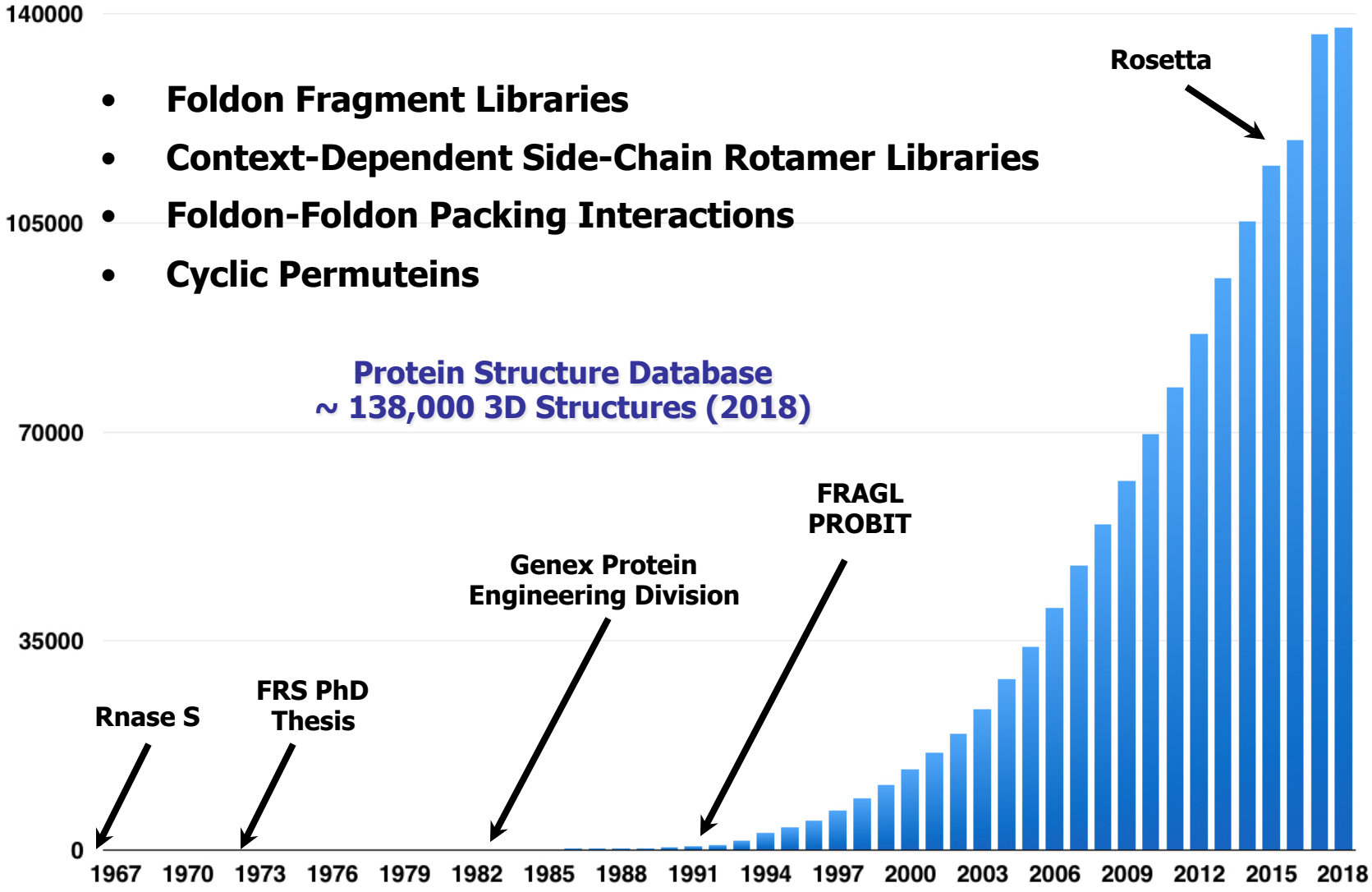
Metamaterial Development Process: A Convergence of Enabling Technologies



IMIPLEX NanoSystems Inc

Heuristic Protein Design

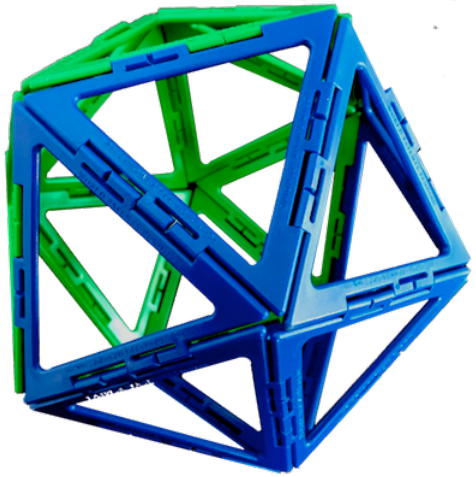
- Foldon Fragment Libraries
- Context-Dependent Side-Chain Rotamer Libraries
- Foldon-Foldon Packing Interactions
- Cyclic Permuteins



IMIPLEX NanoSystems Inc

Nano-Architecture Principles

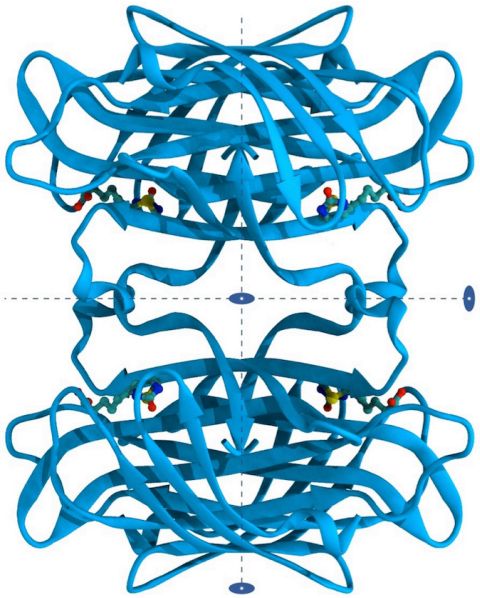
Modular node and strut components connected to form alternative 2D and 3D structures

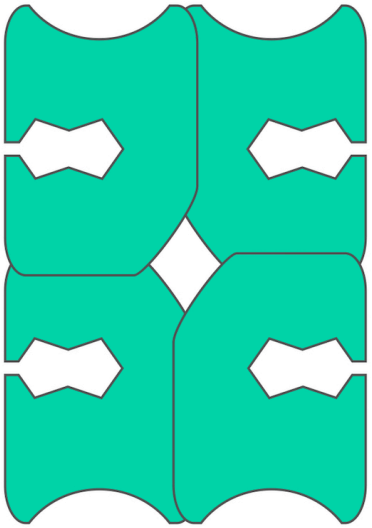


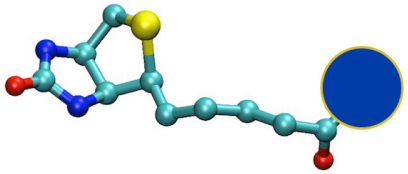
Custom engineered component interfaces

Engineered molecular components that incorporate multiple, alternative, functional subdomains

Streptavidin Struts for Nanostructure Assembly



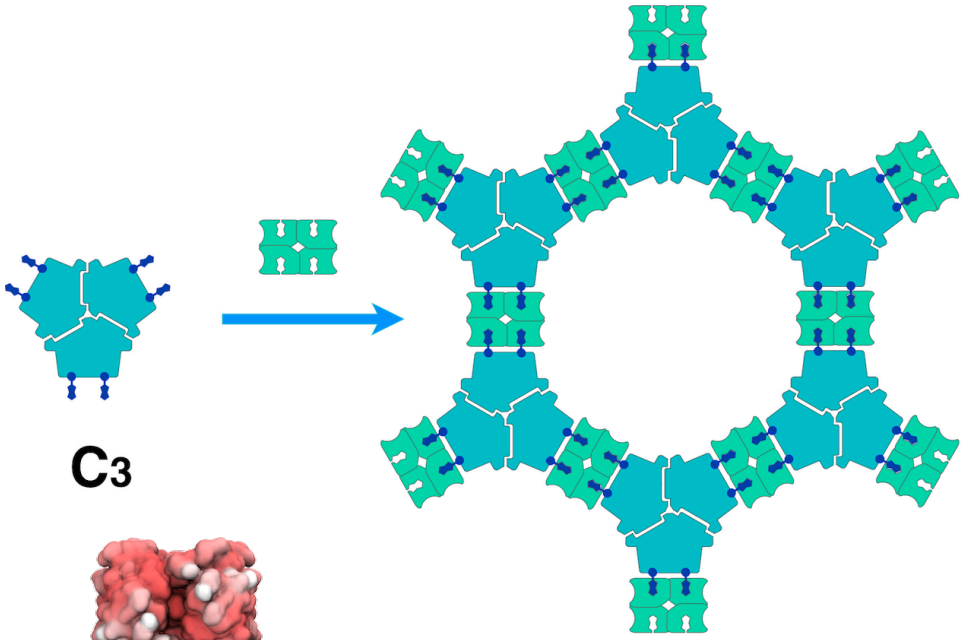




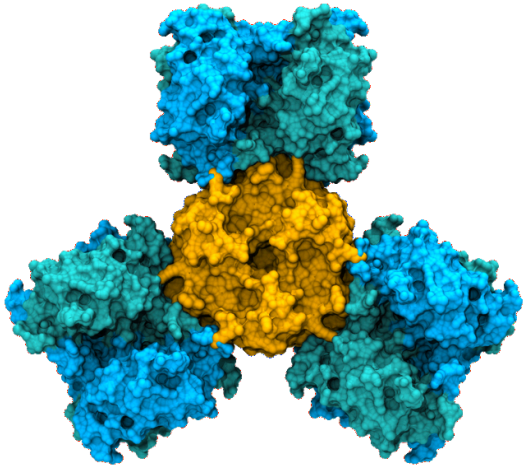
Biotin

Streptavidin D2 Tetramer binds 4 Biotins $K_d=10^{-14}$ M

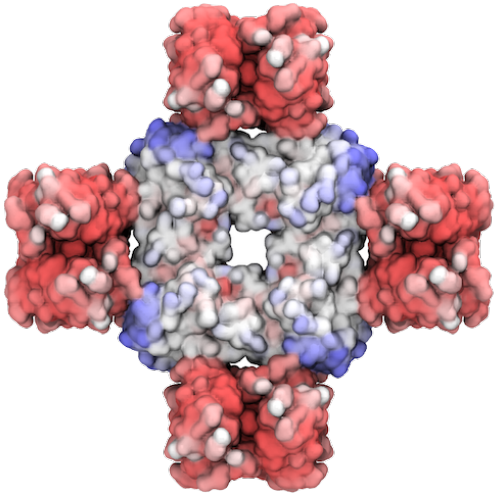
Strut and Node Lattice Assemblies



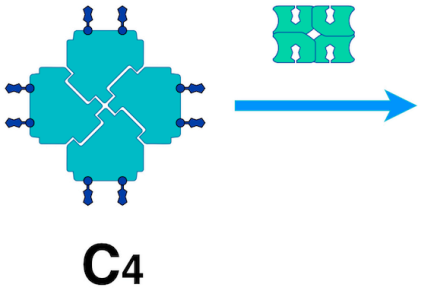
C3



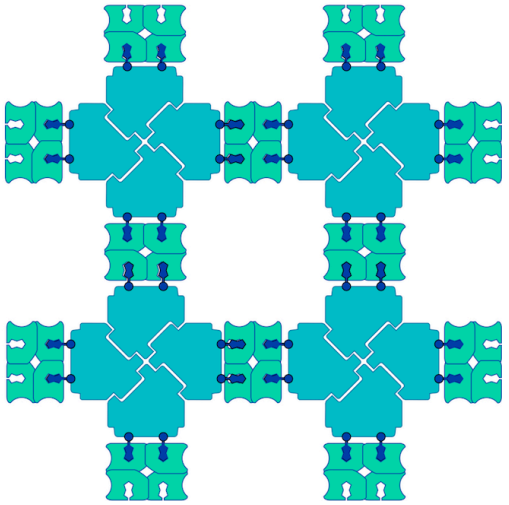
1THJ:3SAV C3



IPP4:4SAV C4

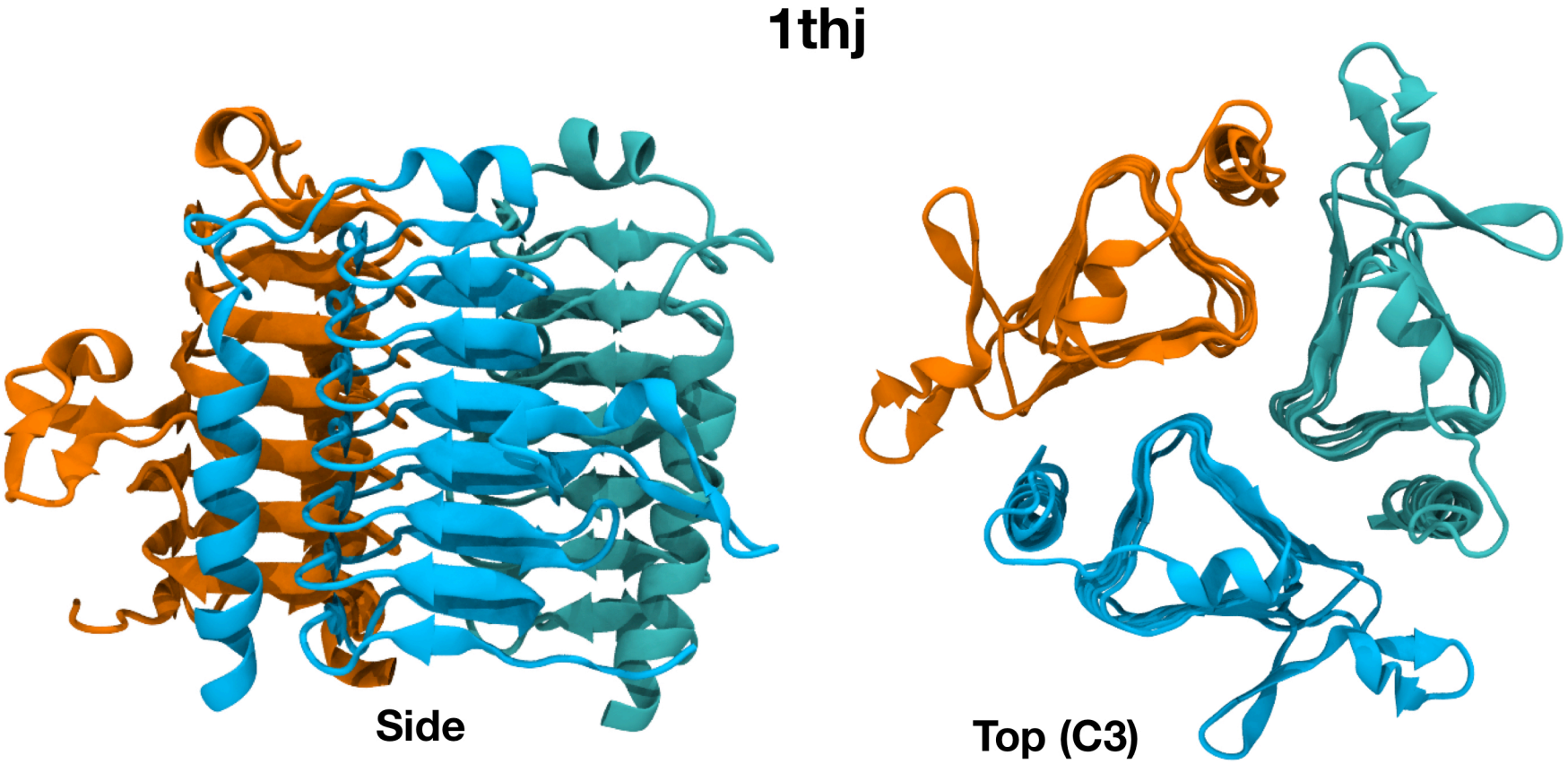


C4



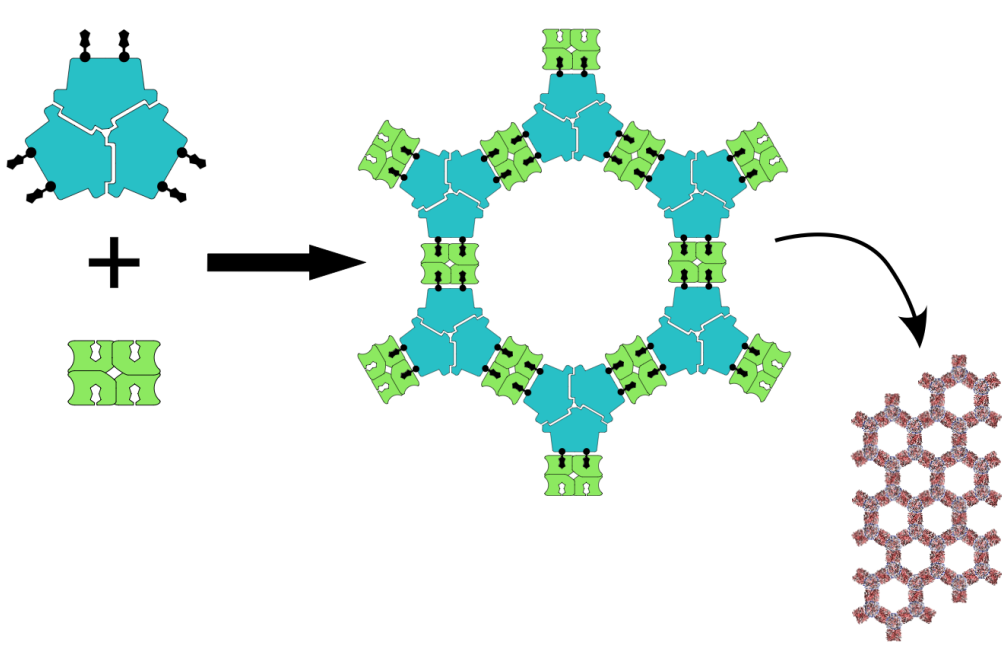
IMIPLEX NanoSystems Inc

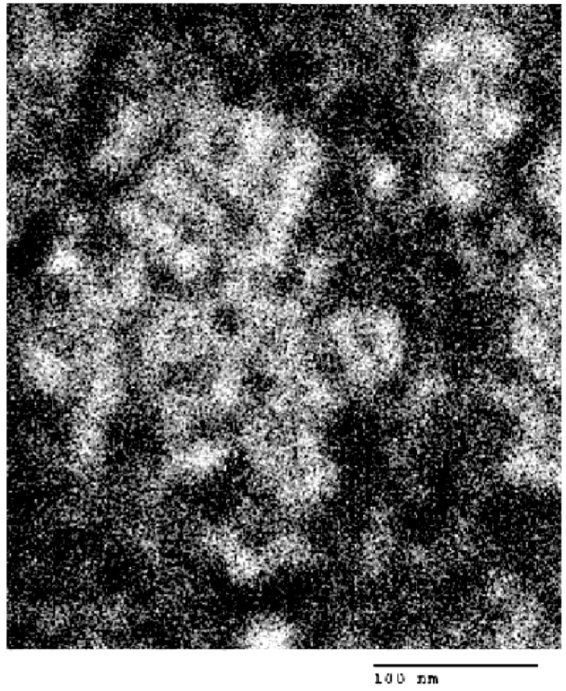
TriPol Node Framework Architecture



IMIPLEX NanoSystems Inc

TriPol Hex Lattice Nanostructure Assembly

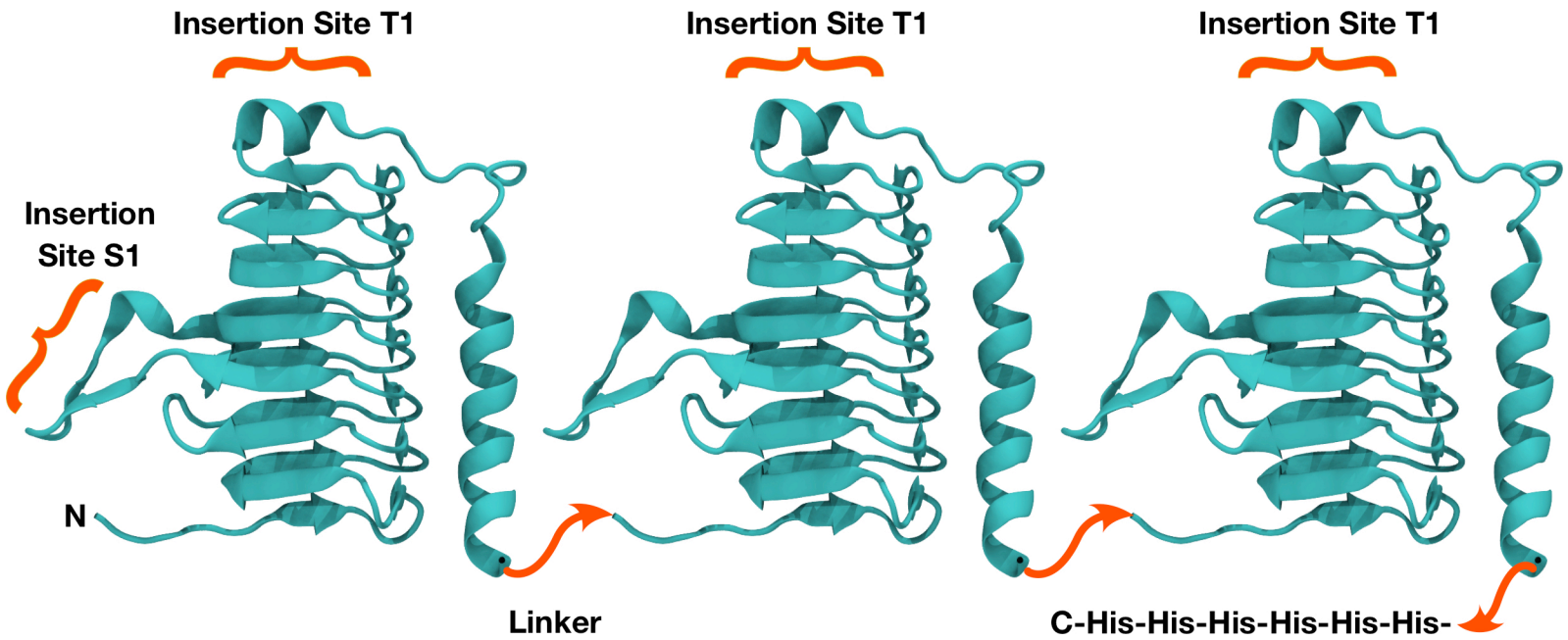




With M. Bisher, Princeton University

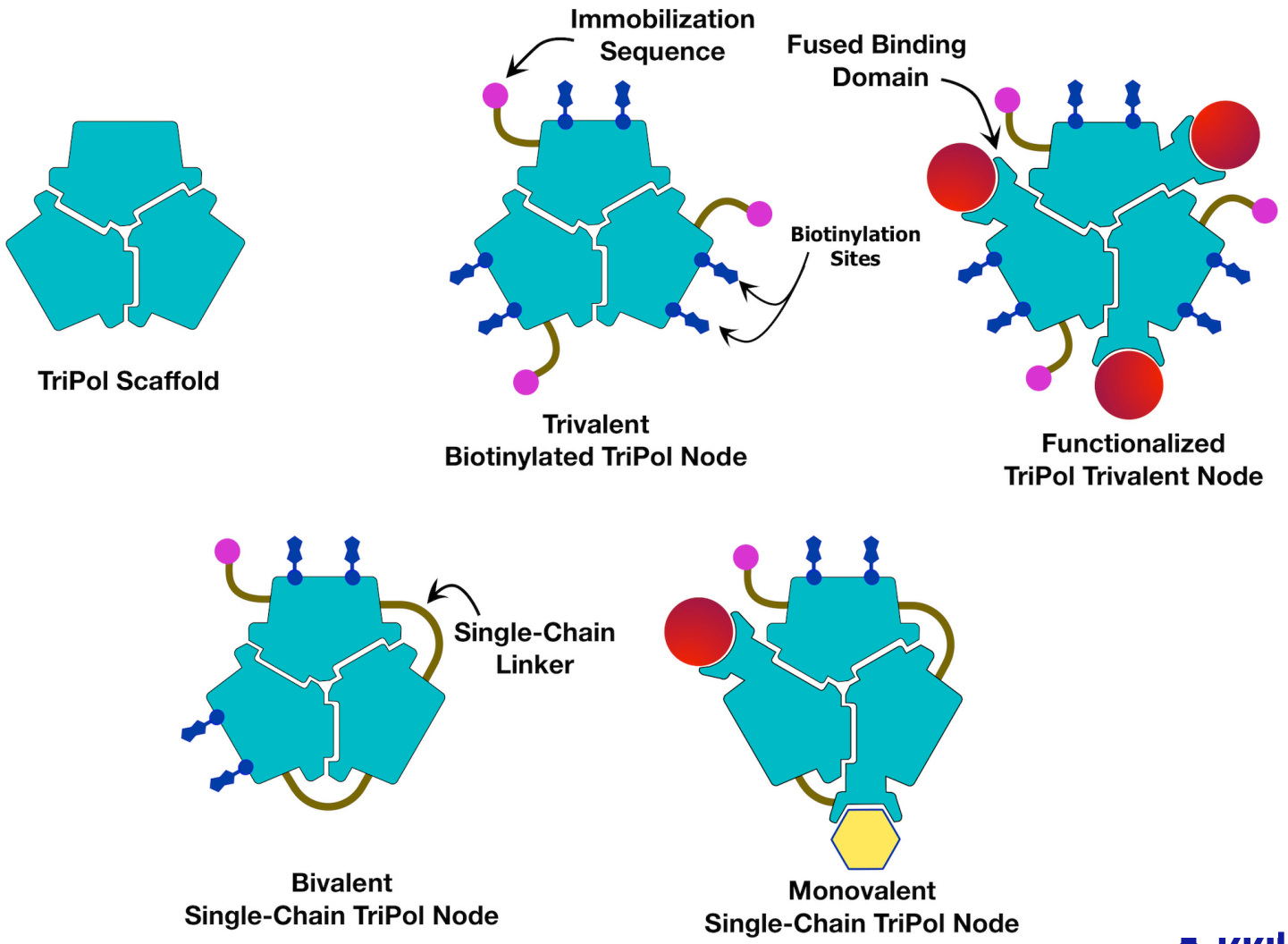
IMIPLEX NanoSystems Inc

TriPol Node Single Chain Construct

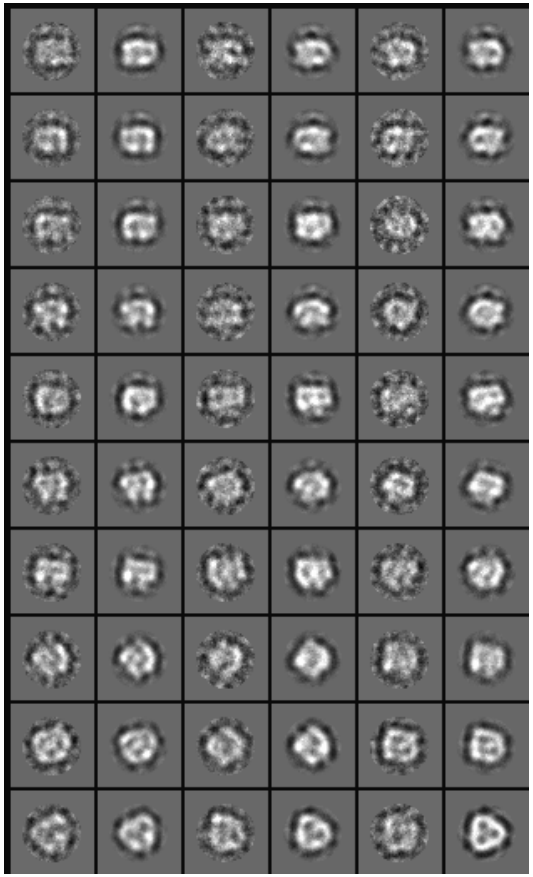


IMIPLEX NanoSystems Inc

TriPol Node Variants

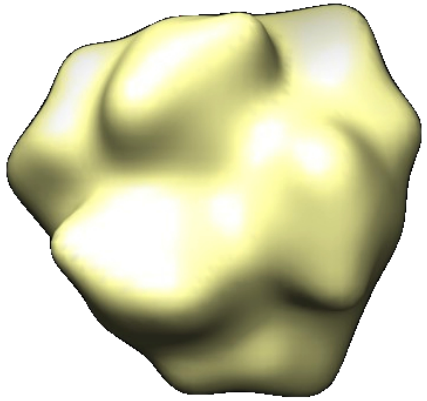


EM Image Reconstruction of TriPol Single-Chain Node

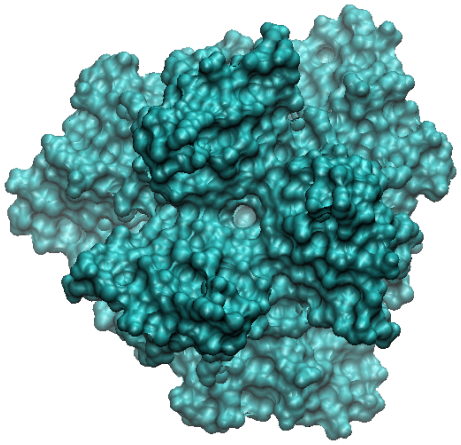


60 EM Images

Schematic



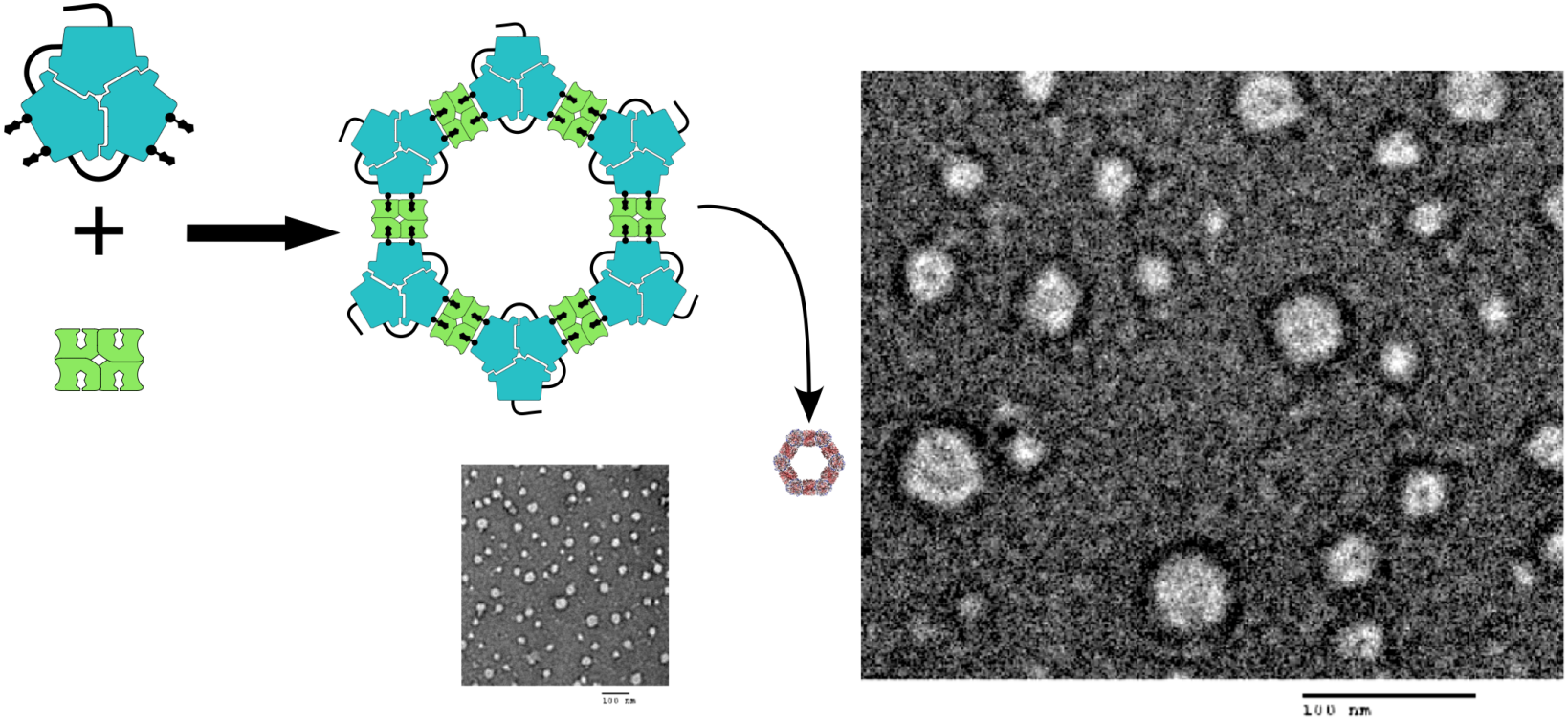
Averaged EM Image



X-Ray Structure Molecular Surface

EM Images and processing by Dr. Rubin Diaz-Avalos (David Stokes Lab) New York Structural Biology Center

Hexagonal TriPol Nanostructure (Edge ~10nm)

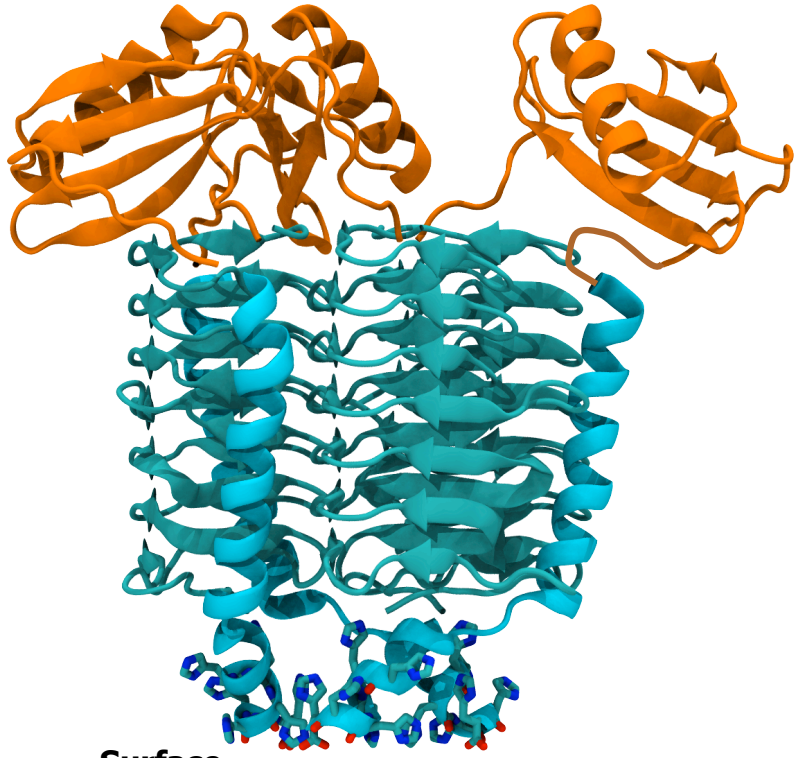


With M. Bisher, Princeton University

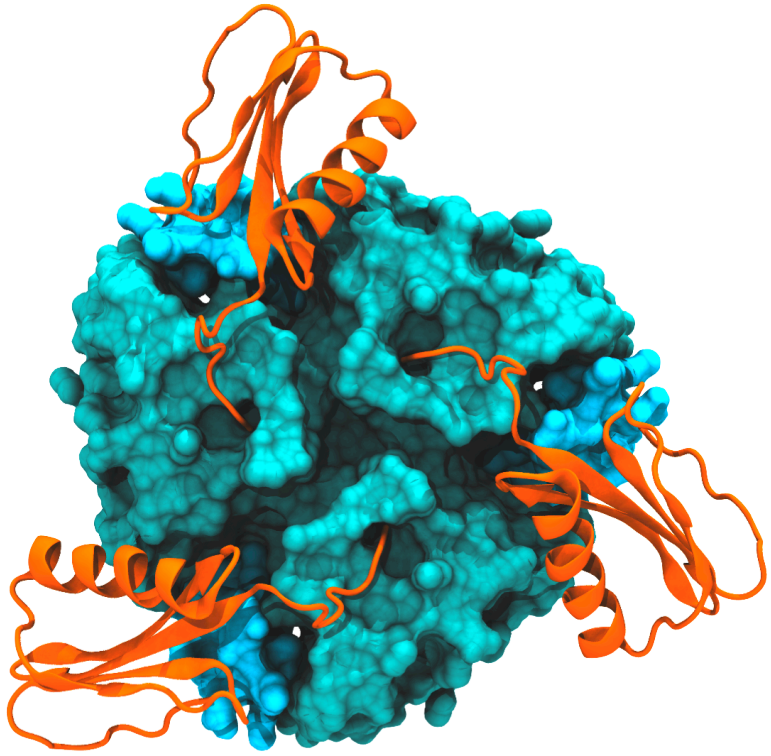
IMIPLEX NanoSystems Inc

TriPol-Ig

Protein G
Domain Fusion



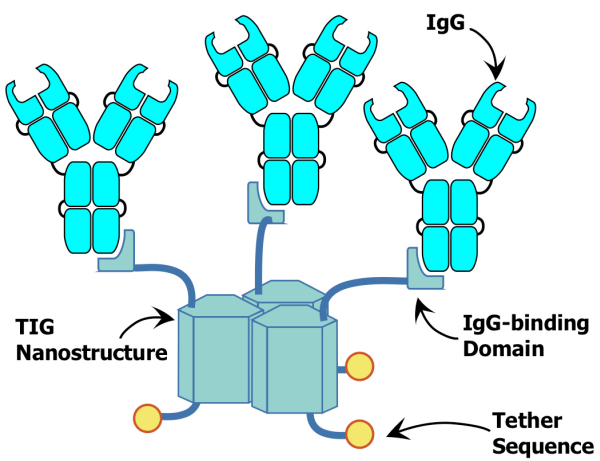
Surface
Immobilization
His Tag



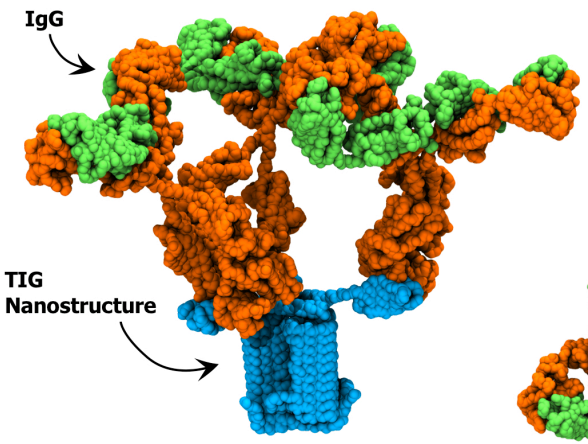
Top View

IMIPLEX NanoSystems Inc

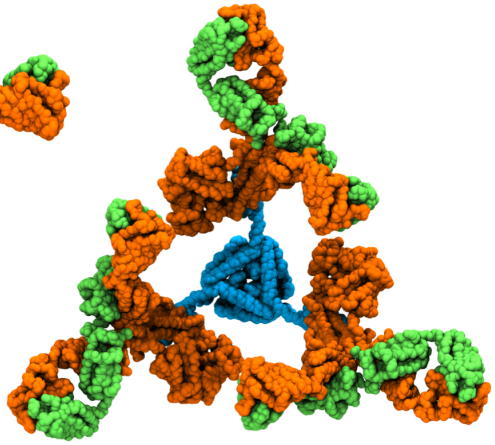
TriPol-IgG Nanostructure



Schematic



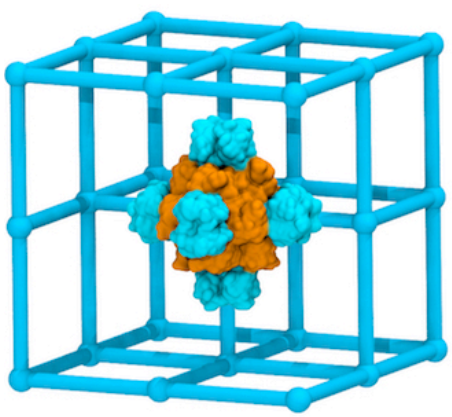
Molecular Model



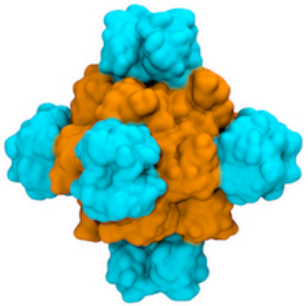
Molecular Model (Top)

- **Plug-and-Play utility in numerous applications that depend on IgG binding for detection specificity**
- **Improve reproducibility and introduce avidity-based binding enhancement owing to geometrical control over antibody presentation**

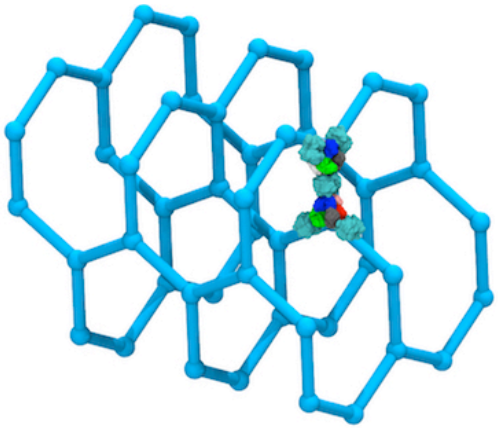
Strut & Nodes Molecular Architecture (3D)



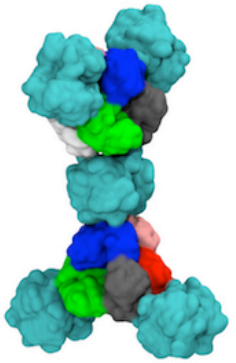
Cubic Lattice



Cubic Node



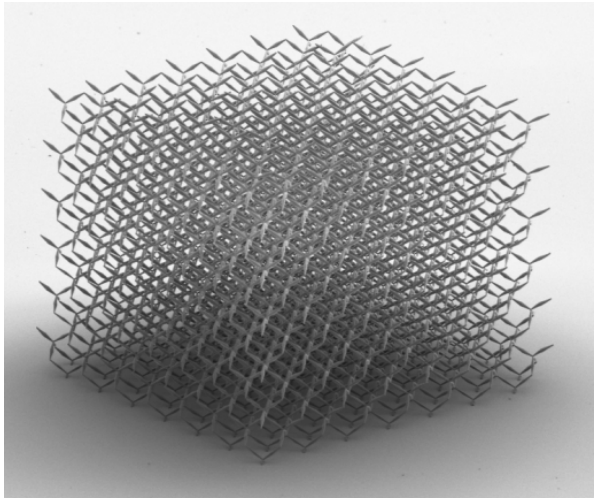
Trigonal Lattice



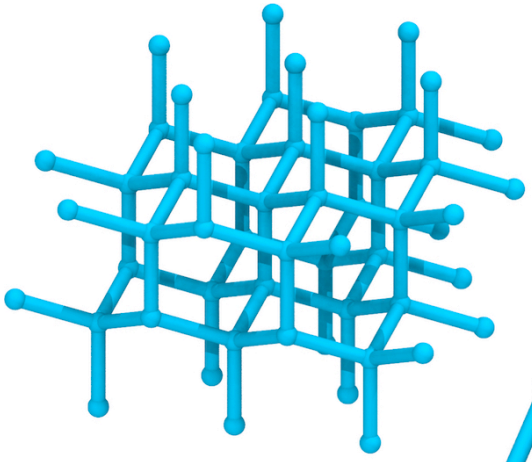
Trigonal Node

- **Multimeric protein nodes with 3D point group symmetry can be functionalized with biotin groups and interconnected with streptavadin tetramers to form 3D structures and lattices with defined molecular geometry.**

Structural Metamaterials

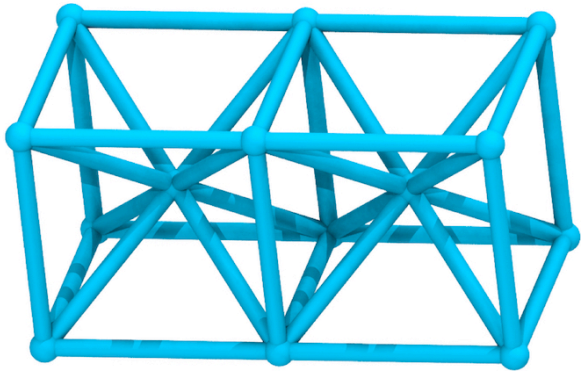


Pentamode Metamaterial (finite compressibility, near-zero shear modulus)



Diamond Lattice

Reinforced Cubic Lattice



- **Nanoscale 3D lattices can form the basis for advanced materials with strength/weight and viscoelastic properties not found in conventional materials**

Metamaterial Photonic Structures

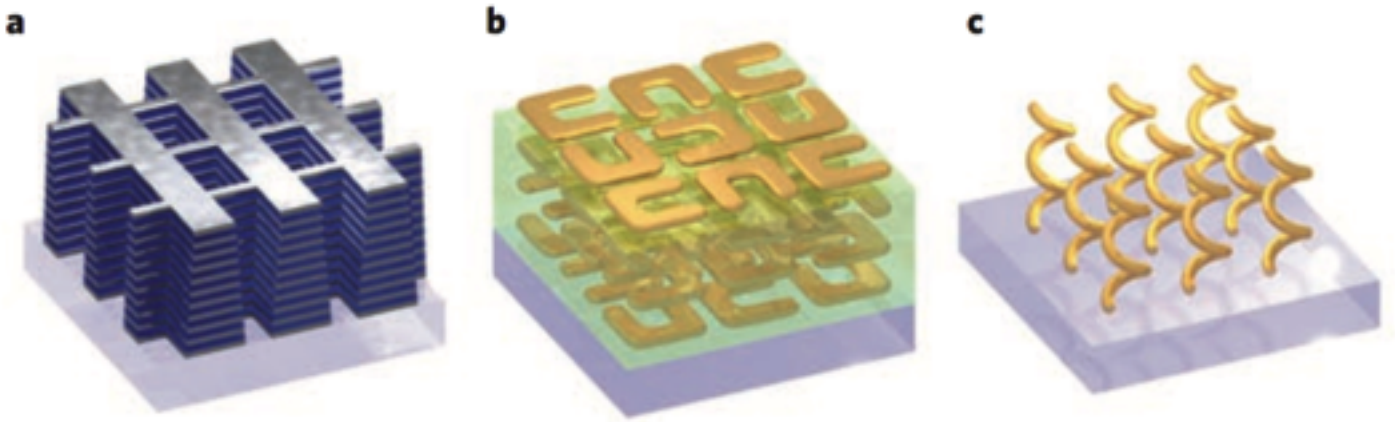
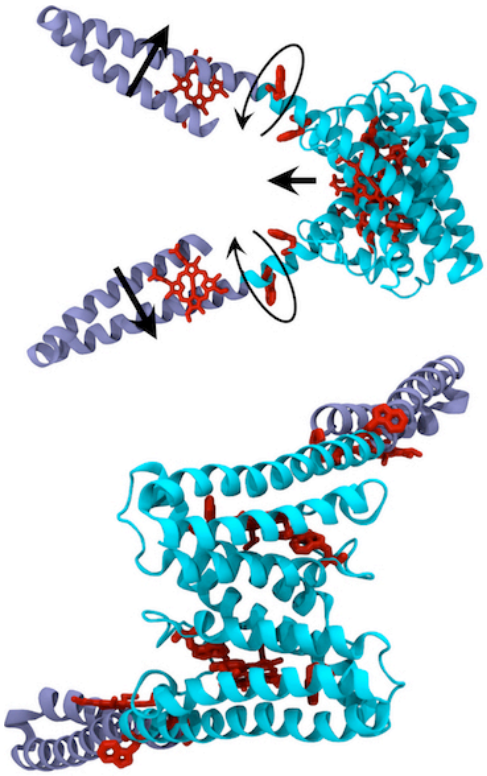


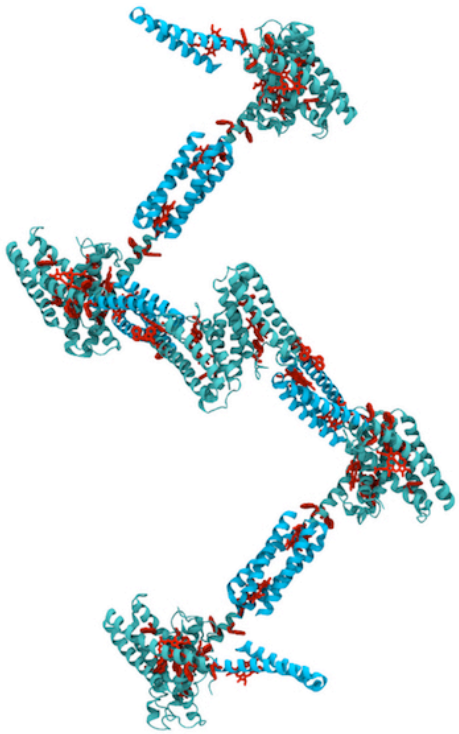
Figure modified from Soukoulis and Wegener *Nature Photonics* (2011)

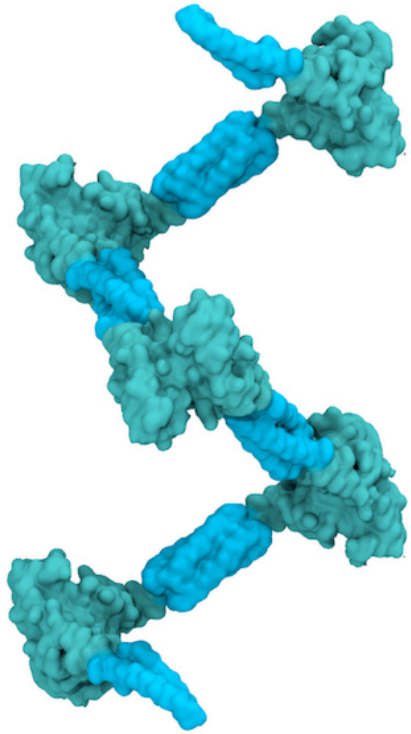
- **Inorganic photonic metamaterials derive function from the precise size, shape, and spatial organization of their constituent repetitive patterns, organized at scales that are smaller than the wavelengths of the phenomena they influence**
- **Metamaterials can block, absorb, enhance, or bend electromagnetic radiation in ways that are not possible using conventional materials**
- **Novel applications include cloaking devices, superlenses that are not limited by conventional diffraction limits, novel types of antennas, and additional optoelectronic applications**

Helical Nanostructure for Photonic Applications



Adjustable Dimer Module

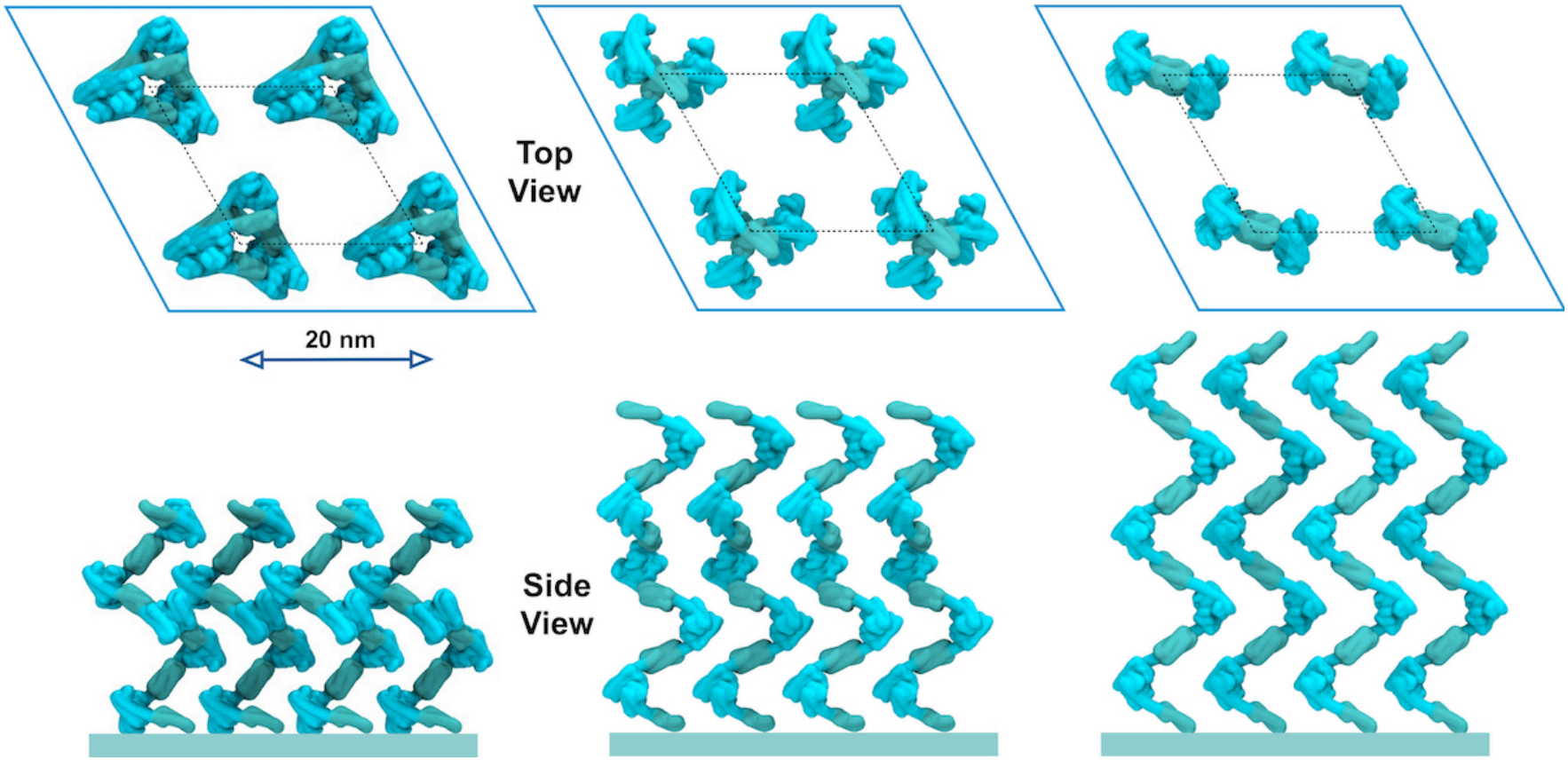




Helix Assembly

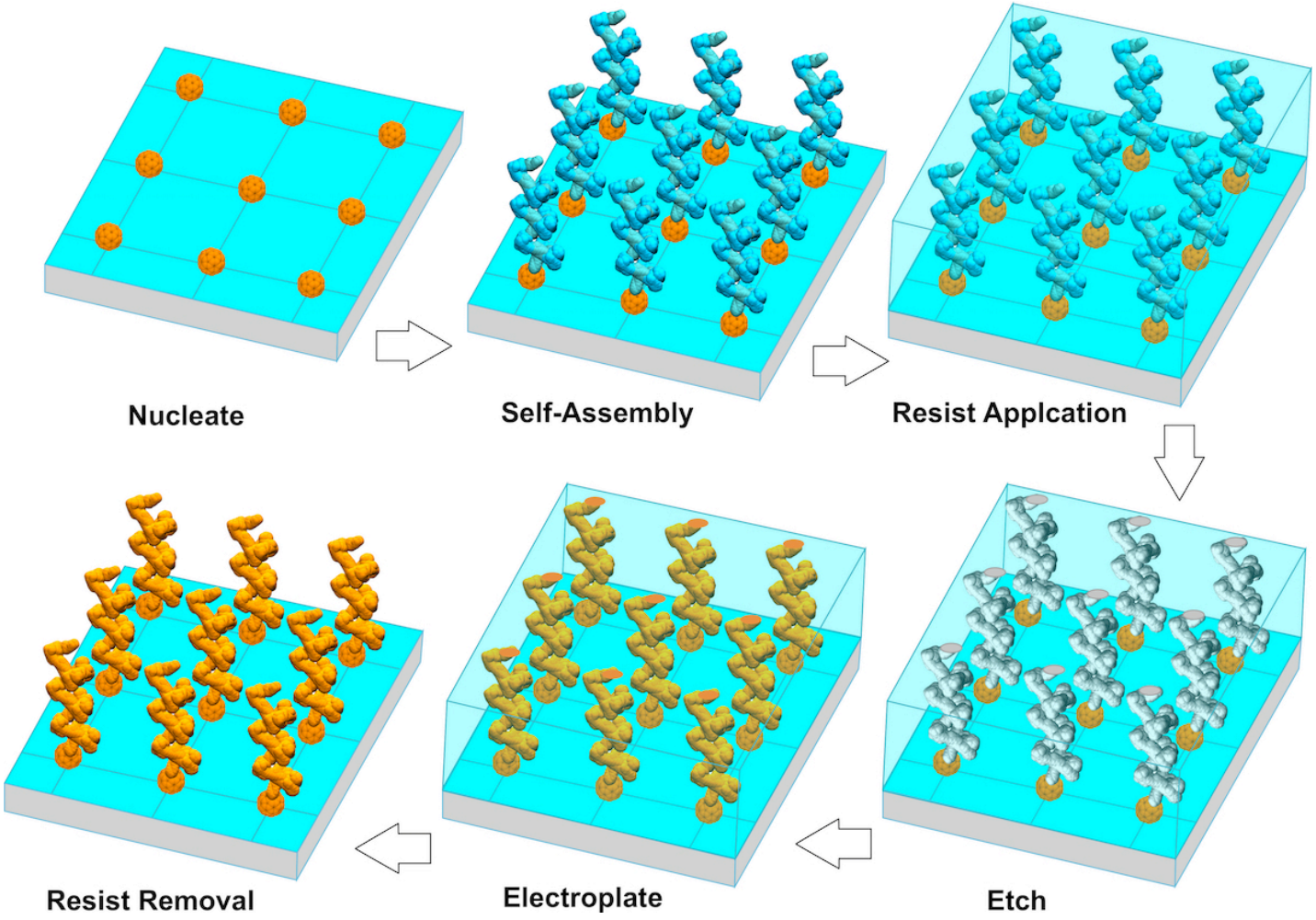
- **Engineered protein:protein interactions provide an alternative strategy for building protein nanostructures that spontaneously assemble.**

Photonic Helical Metamaterials



- **Geometry of helical assemblies can be controlled through protein engineering of helical dimer components**

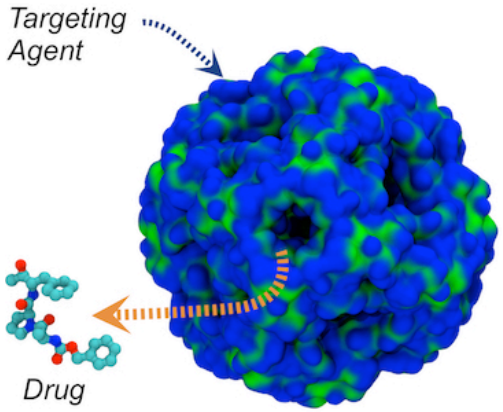
Inorganic Metamaterial Fabrication Process



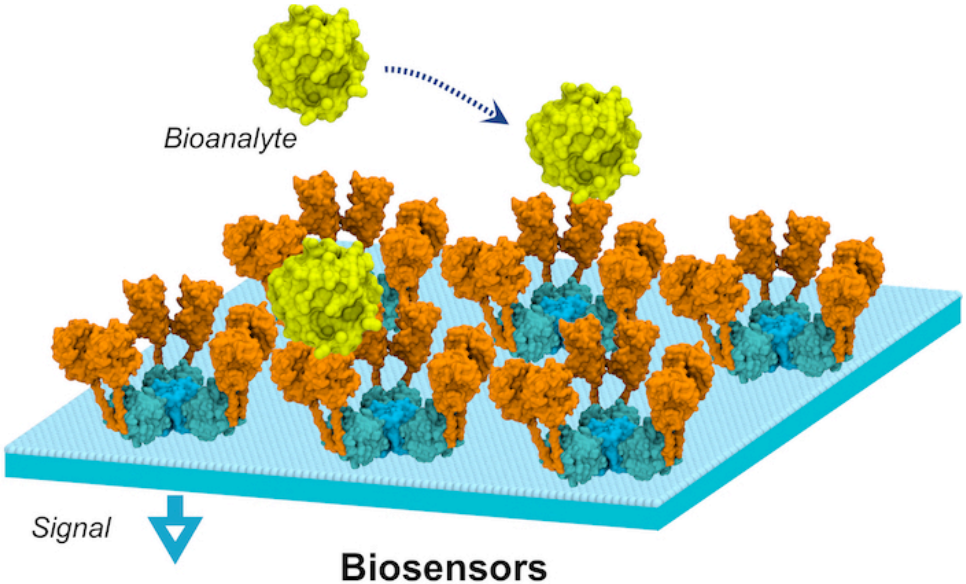
- **Self-assembled protein nanostructures provide patterns for nanoscale 3D resists**

IMIPLEX NanoSystems Inc

Organic Metamaterials: Biomedical Applications



Targeted Drug Delivery



Biosensors

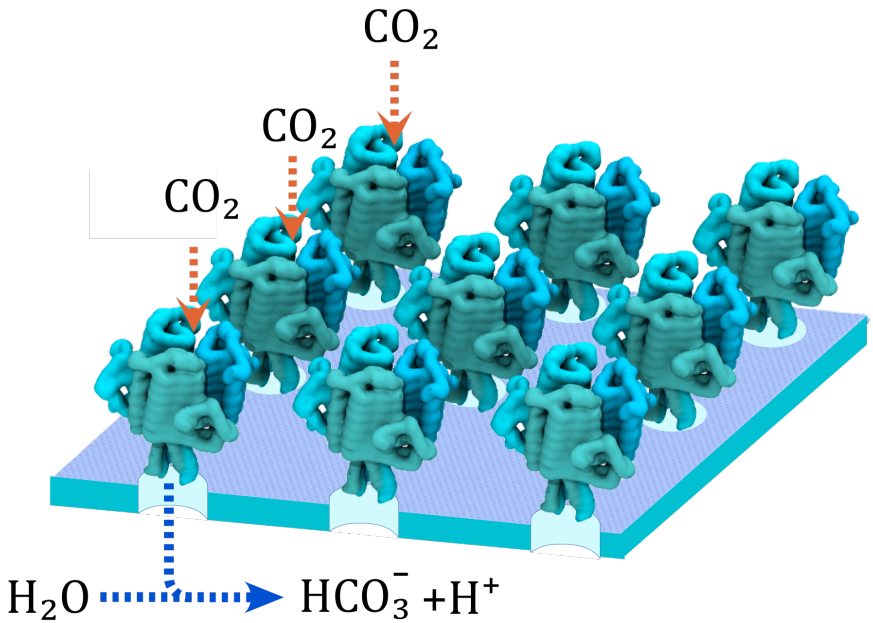
- **Active biomaterials opportunities span a wide range of applications**
- **Medical applications facilitated owing to intrinsic compatibility between living systems and protein-based nanostructures**

IMIPLEX NanoSystems Inc

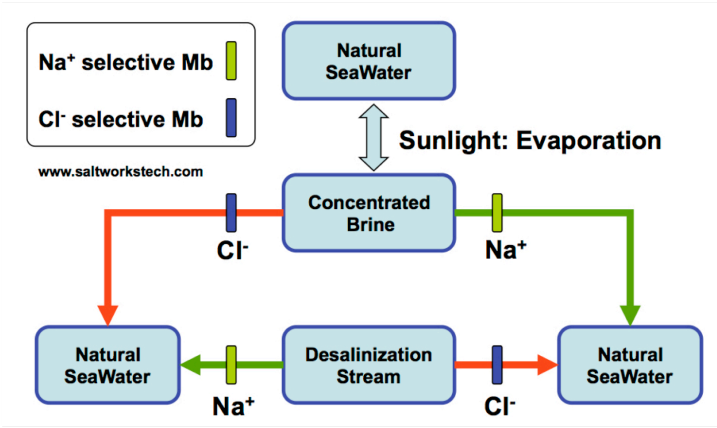
Organic Metamaterials: Vectorial Chemistry

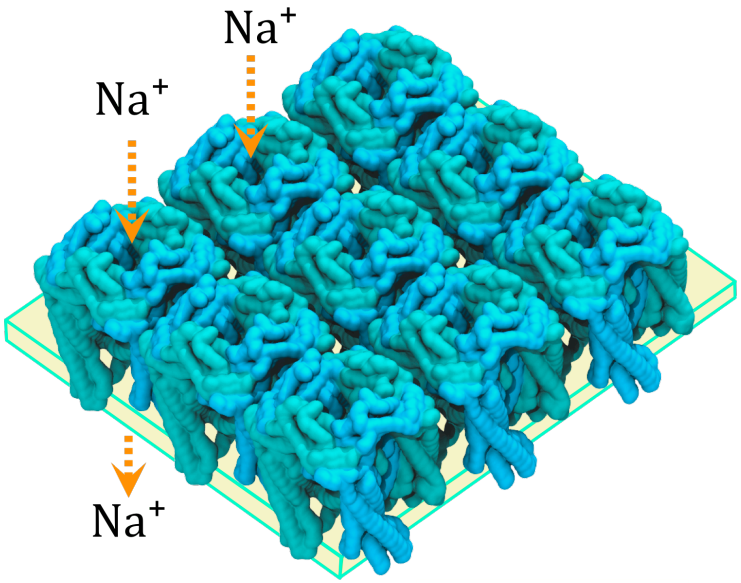
Carbon Fixation

$$H_2O + CO_2 \underset{}{\overset{\text{Carbonic Anhydrase}}{\rightleftharpoons}} HCO_3^- + H^+$$



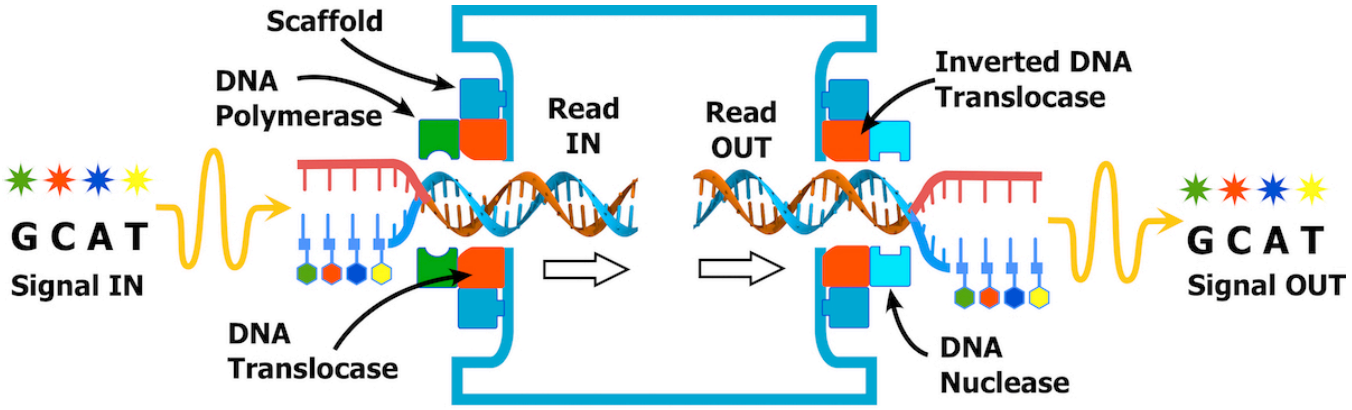
Desalinization



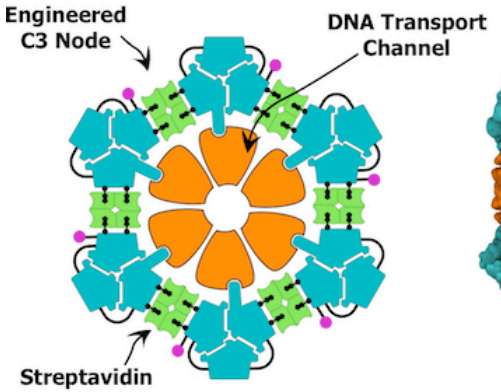


IMIPLEX NanoSystems Inc

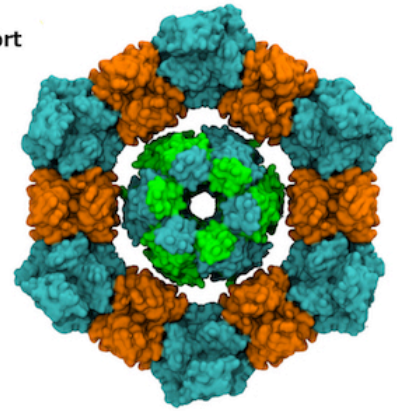
Information Technology: Molecular Data Storage



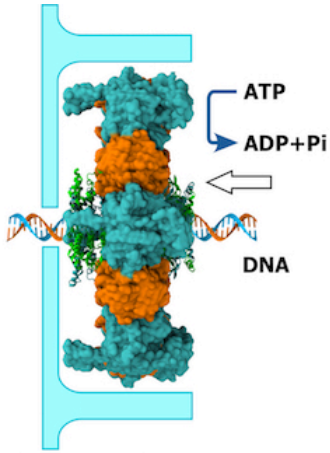
DNA Memory Schematic



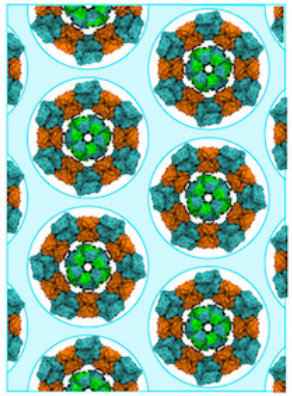
Schematic



Molecular Model



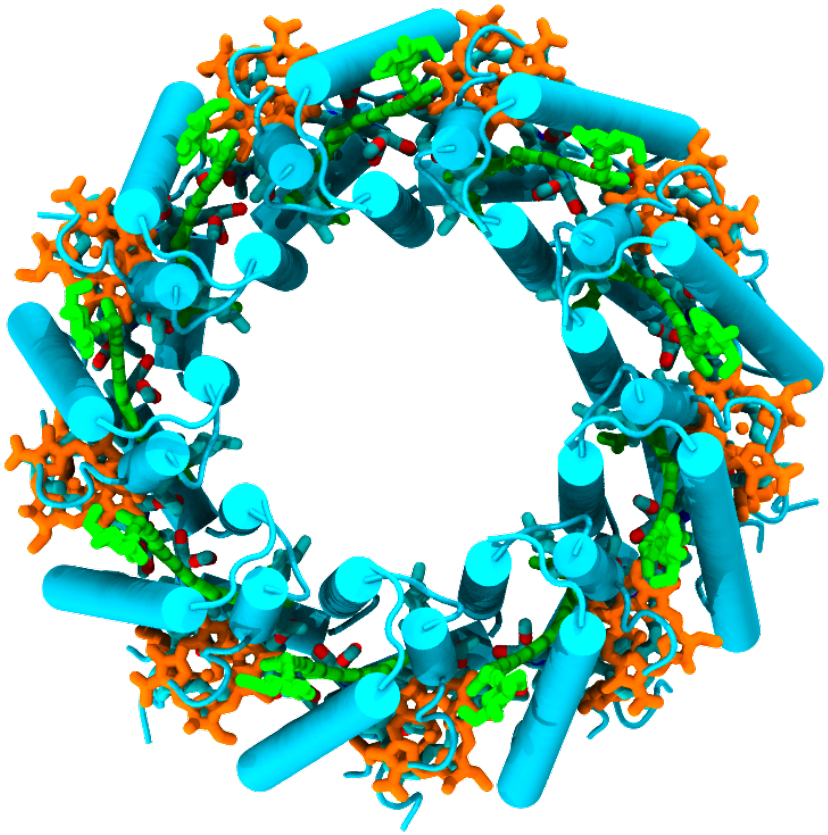
Device Integration



Multichannel Substrate

IMIPLEX NanoSystems Inc

Organic Molecular Electronics:
Optimal function requires precise control intermolecular geometry



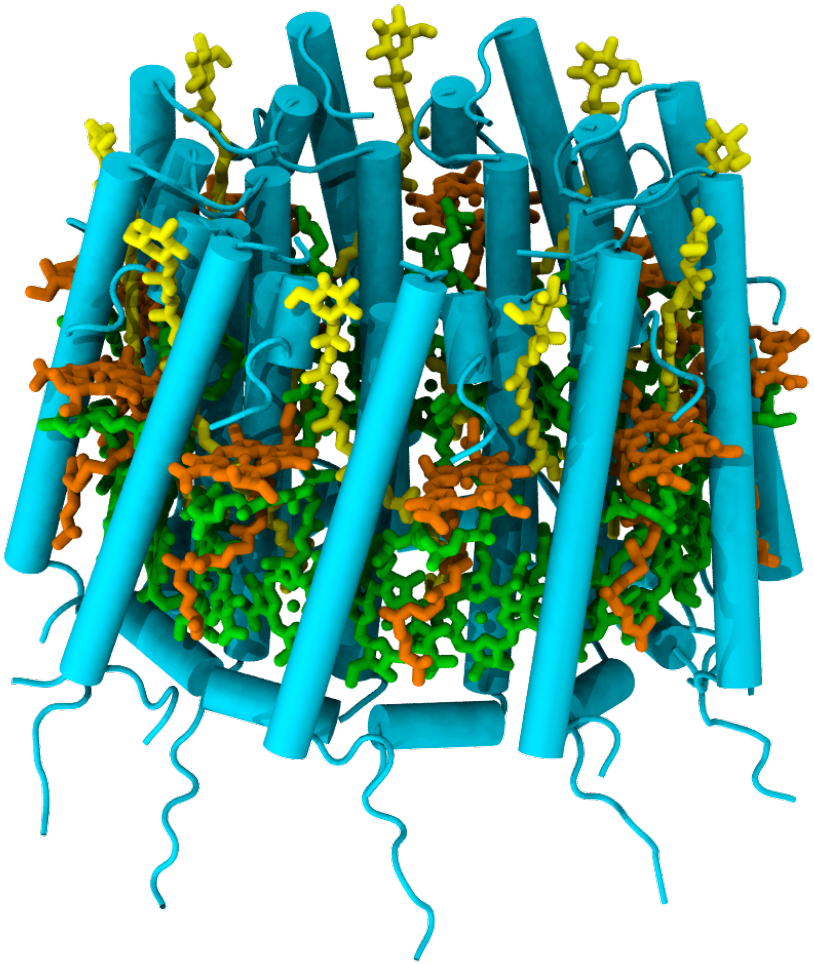
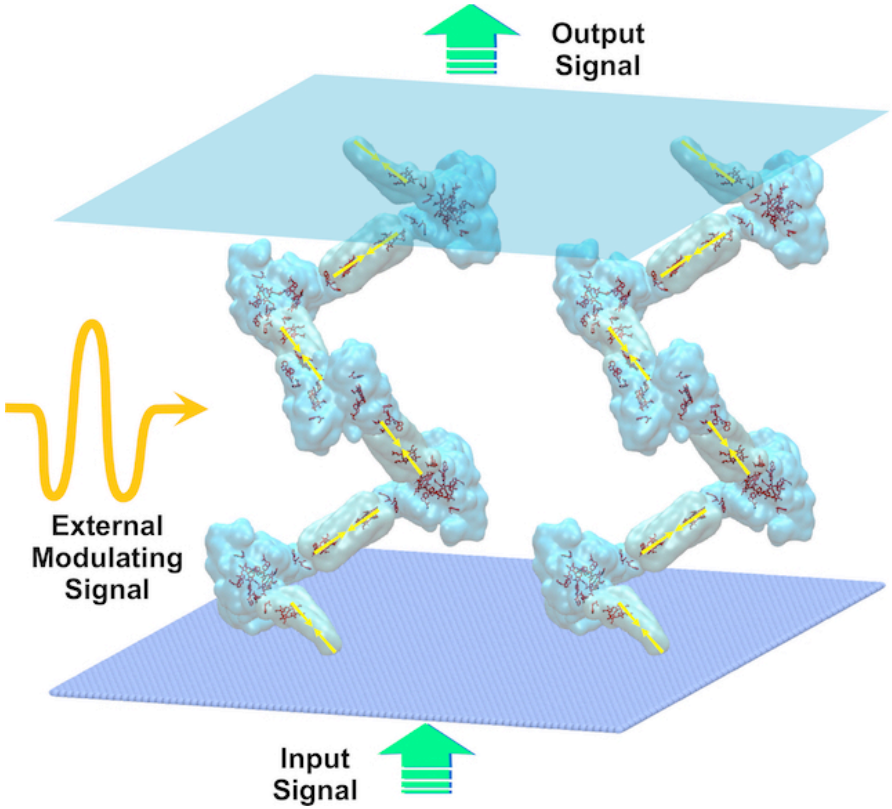
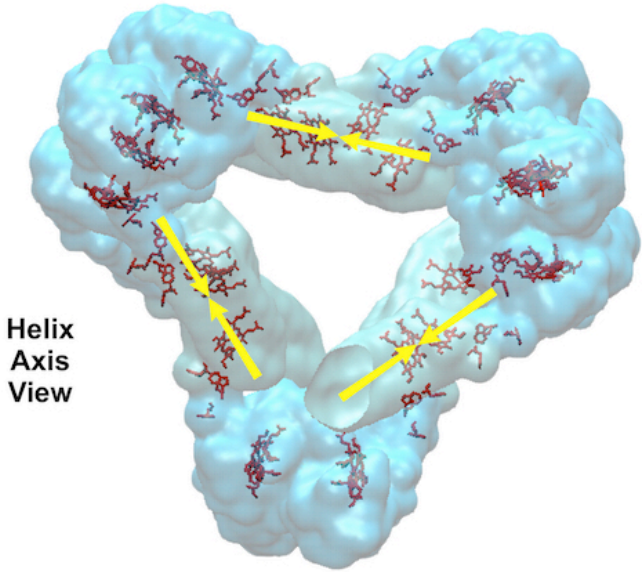


Photo Receptor Antenna LH2
(9 carotenoids, 9 Bclh 800, 18 Bchl 850)

Information Technology: Optoelectronic Systems

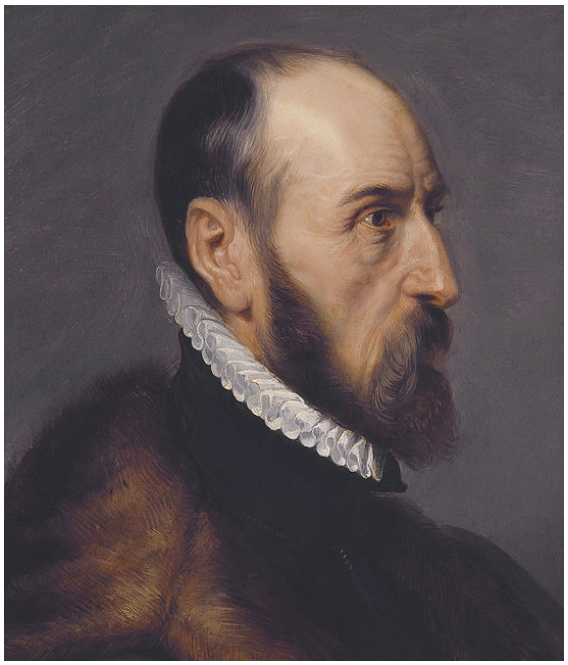




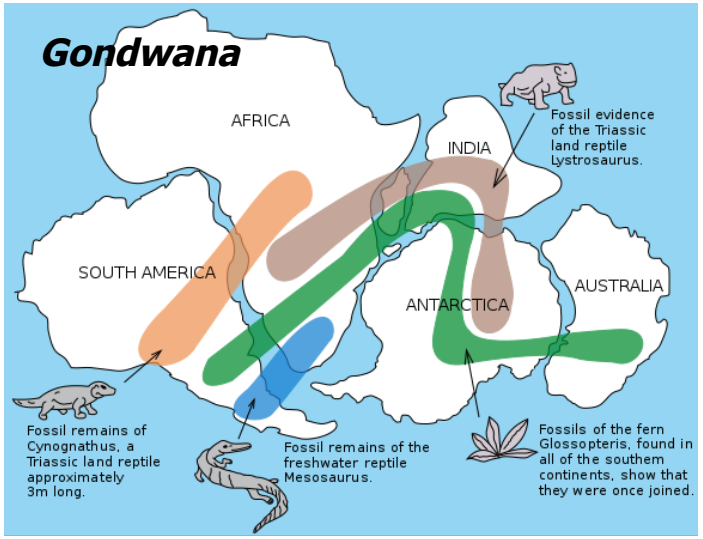
- **Molecular Electronics**
- **Optically transparent nanoscaffolds for nonlinear optical effects**
- **Numerous examples from nature (e.g. avian magnetic cryptochrome navigation)**

IMIPLEX NanoSystems Inc

Abraham Ortelius 1570



Alfred Wegener 1912



Summary

- **Still early days, but technology convergence facilitates the emergence of a nanotechnlogy based on protein self-assembly.**

"Scientists discover the world that exists; engineers create the world that never was."
Theodore von Karman

frs@beta-sheet.org

www.beta-sheet.org