Elements of Biophysics

Saul Pierotti

November 5, 2019

Lecture I and II - General chemistry

- Unit of measure for dipole: Debye (D)
  - $D = 10^{-18} staC * cm$
- Water dipole moment: 1.85 D
- Peptide bond dipole moment: 3.5 D
- Non-bonding interactions: charge-charge>charge-dipole>diploe-dipole>charge-nonpolar>dipole-nonpolar>nonpolar-nonpolar
- Hydrophobiic effect: the tendency of hydrophobic molecules to clump up together in a polar solvent in a way that minimizes the surface exposed to the solvent
  - It is dominated by the entropic component
- Chemical bond: interaction between atoms, moleculs, and ions that makes it possible for them to associate toghether. It is important beacause everithing exists because of chemical bonds. Its potential energy is lower than that of the single elements.
- Intramolecular bonds: metallic, covalent, ionic
- Membrane stability: interaction of polar groups with water, London forces among lipid tails, hydrophobic effect of the lipid tails
- Covalent bond: -40/-110 kcal/mol, 2 Å
- Ionic bond: -5/-10 kcal/mol
- Hydrogen bond: -5/-3 kcal/mol
- Hydrophobic: -1 kcal/mol
- Van der Waals: -0.5 kcal/mol
- Chemical reaction: a process in which one or more reactants are converted to products
- Dissociation process: She means ionization. Henderson Hasselbalch equation
- Molecule: charachterized by its electron density.
- Atomic reactivity: number of electrons in the outermost shell
- Model: equation that describes a phenomenon

Lecture III - DNA

- Human proteins: >100000
- DNA: deoxyribonucleic acid. She wants structure
- Why DNA stores information (ilarious): the univocal base-paring system allows the replication of the information content
- Genetic code: set of rules by which information stored in the genome is translated into proteins in living cells

Lecture IV and V - Thermodynamics

- $\Delta G^0 = -RT \ln K_{eq}$
- $U = Q - W$

- $H = U + pV$
- $\Delta S = \Delta Q_{rev}/T$
- Standard condition: 298 K, 1 atm, pH 7, all reactants 1 M
- Variables: a term in an equation that is free to change independently, or in relation to another variable
- Parameters: a term in an equation that defines the relationship between variables
- Chemical potential: molar chemical energy of a substance $\mu = G/N$
- Kinetics: adds time as a new variable
- Enzyme kinetics: Michaelis-Menten ($V = \frac{V_{max}*[S]}{K_m+[S]}$) and Hill ($V = V_{max}\frac{([S]/[S]_{0.5})^h}{1+([S]/[S]_{0.5})^h}$)
- Kinetic constant: proportionality factor that relates the concentraion of substrates and the reaction rate
- $K_eq = \frac{k_{fw}}{k_{rv}}$

Lecture VI - Proteins

- Protein: a biopolymer composed of aminoacid residues joined toghether by repetitive bonds
- Proteomics: study of the whole protein content in a specific cell type.
- Secondary structure: $\alpha - helix$, $\beta - sheet$, turns
- Protein function: for enzymes it can be defined on the basis of the reaction that they catalyses. For non-enzymes GO-terms
- Reactome: sum of biochemical reactions in a cell type
- When do I know the structure of a protein: when I know its electron density
- Biological process: series of biochemical reactions catalysed by enzymes.
- Metabolic pathway: biological process that converts reactants in products.
- Why -omics: they reflect an olistic view of biology. An interpretation of the cell as a netweork of interaction entities with emergent proprieties, that cannot be easily inferred from the single components
- PDB: 157000 structures

21/10/19

- The course is made of elements, it is not a complete course
- We want to understand biomolecules
- Structure-function relationship
- The scale of things
- I don't know something if I cannot describe it with a model
  - A model is an equation
- Laser time dependent X ray cristallography allows to see a molecule "breathing"
- The deeper you go in resolution, the more you have to increase the perturbation
  - Bragg diffraction theorem
- Life makes sense only in the light of thermodynamics
- A cell is held together by London forces
- Freeze fracturing
- Atomic force microscope
- rita.casadio@unibo.it
- Complexity means that there are emergent proprieties
  - We do not have models that can predict a cell from its components
- Mitosis, meiosis, apoptosis, developmental biology
- Protein synthesis
- In a cell there are at least 240k different proteins
  - Their relative concentration is of paramount importance
- We are able to label all the neurons in a mouse brain

22/10/19

- I came late
- Bonds can be polar
- Ionic bonds are 2.7/3Å long, and around -5 to -10 Kcal/mol
- Covalent bonds are around 100 Kcal/mol and they show an optimal nuclear distance, which is the bond lenght $(1 - 1.4Å)$
    - The typical C-C bond is 1.4Å
- Bond lenghts are calculated by X-ray diffraction studies, bond strenghts with calorimetry
- Disulphide bond are around -40 Kcal/mol, 2.07Å
    - This bond is stable in an oxydized ambient
- Redox ambient potential describes the tendency of molecules in an environment to lose or acquire electrons
- Steric hindrance
- Bond resonance
- Peptide bond is around 1.5Å (?) and is polar,
    - A long protein chain is highly polar end to end 2-4 Kcal/mol
- A Debye is the unit of measure of permanent dipoles
- A water molecule has a dipole
- The C-H bond is not polar (!)
- The OH_3^+ structure is called hydronium ion
    - One O atom generally interacts with 4 H, binding 2 at a time in a covalent way
- An H bond can be completely explained only with quantomechanics, it is 3-5 Kcal/mol and 1.5Å
- Cell membrames

23/10/19

- The dipolar interaction energy can be derived from the dipol moments
- The London force is the most dependent on distance
- The hydrophobic effect is the tendency of nonpolar molecules to aggregate in a polar solvent so to minimize the surface exposed to the solvent
- DNA
- We do not know the structure of mRNA
- The width of the DNA helix depends on salt concentration
    - The more salt the more compact it is
- The information content of DNA is in AT CG base pairing (so she wants)
- Transcribed regions are called loci
- Trascription is controlled by DNA methylation

24/10/19

- The relative dielectric constant comes from condensers
- A biological process is a series of chemical reactions catalized by enzymes
- Protein biosynthesis is a biological process that transfer information from DNA to ribosomes
- The receiver of information is the ribosome
- Every protein has an half-life: there is turnover
- The genetic code is redundant

28/10/19

- A protein is a biopolymer made of repeating peptide bonds, which form the backbone
- When an aminoacid enters in the structure of a protein, it becomes an amino-residue

- An aminoacid has a carboxil group and an aminic group bonded to a C_α
  - The side chain R can be one of 20 varieties
  - The carboxil group has a $pK_a \approx 4.2$
  - The amino group has a $pK_a \approx 9$
- The formation of the peptide bond is a condensation process where a water molecule is expelled
- The peptide bond is planar and has a dipole moment of around x debye
  - The bond is subjected to resonance because the C_α is bound to 2 electronegative atoms, O and N
  - It is shorter than a single bond, and the bond with O is longer than a double bond
  - The plane is defined by O, C_α and N
- Can every sequence be a protein?
  - A protein is long between 50 and 200-300 residues
  - If it is shorter usually it cannot fold in solution
- The bond between C_α and N is called Φ, the bond between C_α and COOH is called Ψ
  - Rotation is possible around these bonds
- Amino acids are zwitterionic, meaning that they have at phisiological pH both positive and negative charge
- The protein structure is the golden standard of proteomics
  - It allows the best predictions on function
- Functional genomics is the use of available data to infer information about unknown genes and proteins
- The secondary structure derives from the stabilisation of the protein by hydrogen bonding
- SCOP categorizes protein domains
  - The class of a protein is the way in which the backbone is organised (secondary structure)
    * An all-alpha or all-beta protein have more than 90% of the residues in that conformation
    * Alpha+beta proteins have a linker connecting the 2 domains
    * Alpha/beta proteins have a mixing of the structures
  - A fold is the topological arrangment of a portion of secondary structure
  - A superfamily has a probable common ancestor
  - A family is a collection of proteins that have similar function and structure
    * They can also have very different sequences (!)
- To categorize proteins, we have to do a structural alignment
- Proteins are a frustrated system because mainly hydrophobic
  - They tend to fold in solution
  - The ΔG associated with the folding process is around 0-50 kcal/mol and it is mostly indipendent from the protein itself

29/10/19

- A protein is stable because it is capable of auto-organization in a polar solvent, which gives rise to million of pairwise interaction
- Proteins form crystals when the solution is super-saturated
- Proteins are social entities
- The folding of a protein is defined hydrophobic collapse
- Chaperons help the folding process in vivo
- Folding in vitro happens in a binary system made of the protein and water
- It is better to say compute than predict
- What is protein function?
  - It can catalize a chemical reaction
- In bioinformatics, enzymes are classified in clusters by the Enzyme commission classification (ECC)
- The EC classification is a hierarchical system
  - EC1 are oxydoreductases
  - EC2 are transferases
    * They transfer a functional group from a molecule to another
  - EC3 are hydrolases

- EC4 are lyases
  - * A lyase is an enzyme that catalyzes the breaking of a covalent bond by a mean different from hydrolysis
- EC5 are isomerases
  - * They promote intra-molecular rearrangement
- EC6 are ligases
  - * A ligase forms new bonds using ATP
- EC7 are translocases
  - * It is a fancy name for membrane transporter
- Numbers after the first narrow down the activity of the enzyme
- The EC number does not consider isoforms (!)
  - It cannot differentiate
- GO terms are a big achievement of bioinformatics
- GO means gene ontology
  - An ontology is a vocabulary
- GO terms are standard words or phrases used for indexing or retrieving information
  - They are reviewed by experts in the respective fields
- GO defines relationships between terms
- Cellular component, Biological process and Molecular function are the 3 roots of Gene ontologies
  - From the roots, we can go down a hierarchical organization to find a specific term (es. cytosol)
  - Each term has a random number assigned to it
- The problem of protein validation: we can easily detect putative proteins, but it is difficult to experimentally validate them
- Computational biology is a field of bioinformatics involved in the development of tools for sequence analysis, functional genomics and proteomics
- Systems biology develops models for biological systems
- Thermodynamics dominates our life
  - U is the internal energy of the system, it is composed of a kinetic term and a potential term
  - W is mechanical work on the system
  - Q is heat added to the system
  - Most of the quantities in thermodynamics cannot be measured in an absolute way, only changes are measurable
- The entalpy is the sum of internal energy and mechanical work done on the system
  - $H = U + pV$
  - The mechanical work is composed of the product of pressure and volume
  - A negative change in entalpy means that the process is exotermic
- Entropy
- Gibbs free energy
  - If ΔG is 0 we are in an equilibrium condition
  - In life equilibrium means death
- Kinetics is a branch of thermodynamics

30/10/19

- The equilibrium constant of a reaction $A \rightleftharpoons B$ is $K_{eq} = \frac{[B]}{[A]}$
- ΔG is related to K_{eq} by $\Delta G^0 = -RT \ln(K_{eq}^0)$
- There is a relationship between the kinetic constants and the equilibrium constant
- Enzyme reactions can be modelled by the Michaelis-Menten equation
  - K_m and V_{max} are parameters of the equation
  - The relationship is hyperbolic
  - K_m is the $[S]$ at which $v = \frac{1}{2}V_{max}$
  - The enzymes that follow this kinetic are monomeric
- Multimeric enzymes are modelled by the Hill equation

- The relationship is a sigmoid
- There is an exponential term h
- The parameters are the same of the MM, V_{max} and $[S]_{0.5}$
- This takes into account the allosteric effect