

Energy terms in potential function

This lecture describes the various energy terms in the potential function. Due to additivity principle the potential energy E_p is represented as the sum of local and non-local terms. Local terms include bond-length, bond-angle, dihedral and improper dihedral potentials. Non-local energy terms consist of pairwise electrostatic and van-der-Waals interactions.

I. Bond-length potential

This potential controls the length of covalent bonds. The most commonly used form for bond-length potential V_{BL} is based on the Hook's law (harmonic potential). In this case,

$$V_{BL} = k_a (r - a)^2, \quad (1)$$

where r is the distance between covalently bonded atoms, a is the corresponding equilibrium distance, and k_a is the spring constant. The equilibrium distance between atoms can be obtained from X-ray diffraction experiments or using *ab initio* quantum calculations. The spring constant can be determined using the results of infrared adsorption and Raman spectroscopy. The frequency of bond length fluctuation in Eq. (1) is $\omega^2 = \frac{2k_a}{m}$, where m is the mass of bonded particle. If the bonded masses are comparable (i.e., the motions of both atoms must be taken into account), then the effective mass $\mu = m_1 m_2 / (m_1 + m_2)$ must be used in ω , where m_1 and m_2 are the masses of bonded atoms. Eq. (1) is used in CHARMM or OPLS force fields and the example of Eq. (1) parameterization is given in Fig. 1.

bond (atom types)	k_a , kcal/(mol·Å ²)	a , Å	comment
CT1 C	250.000	1.4900	! backbone C α -C bond
CT1 CT1	222.500	1.5000	! C α -C β bond, e.g., in Val
NH1 H	440.000	0.9970	! backbone N-H bond
CA CA	305.000	1.3750	! C-C bond in aromatic ring
OT HT	450.000	0.9572	! O-H bond in water

Fig. 1 The Hook's law parameters for bond-length potential in CHARMM22.

Bond-length harmonic potential is applicable for small bond length deformations (< 10%). Although at larger deviations the covalent bond breaks, the harmonic term (Eq. (1)) does not capture this effect. It is possible to use a Morse potential to account for bond breakage. The Morse potential is given by

$$V_{BL}^M = D(1 - e^{-S(r-a)})^2,$$

where D and S determine the value of bond-length potential at $r \rightarrow \infty$ (dissociation energy) and the width of the potential well. Although Morse potential accounts for bond breakage, it remains finite at $r \rightarrow 0$. Because at r close to a the Morse potential can be approximated by the harmonic potential, the Morse parameters D and S can be found from the parameters in Eq. (1). The Morse potential is used to provide better representation of bond vibrational frequencies. As an intermediate solution to improve the agreement of bond length frequencies, one can add several terms from the Taylor expansion of the Morse potential to Eq. (1) (for example, cubic or quartic terms). It is important to note that any polynomial of the odd order used for V_{BL} may potentially lead to catastrophic instability as $V_{BL} \rightarrow -\infty$ with $r \rightarrow \infty$.

In practice, bond lengths associated with light atoms (hydrogens) are often constrained to their equilibrium values using SHAKE or RATTLE algorithms (see coming lectures). In this case, the bond-length energy term is not considered in the energy function.

II. Bond-angle potential

Bond-angle potential is designed to reproduce the bond geometry in molecules, which is in turn controlled by hybridization of atomic (electronic) orbitals. For example, sp hybridization allows to the formation of two bonds with the angle between them of 180° . Other hybridizations, such as sp^2 or sp^3 , result in trigonal or tetrahedral arrangements of covalent bonds around the atom (the respective bond angles are 120° or 109.5°). From these considerations the parameters for V_{BA} can be derived. It must be noted that local environment may distort the bond angles. For example, for water molecule, in which oxygen has a sp^3 hybridization, the bond angle is about 105° .

The following forms of V_{BA} are used

$$V_{BA} = k_\theta (\theta - \theta_0)^2 \text{ (harmonic potential)} \quad (2)$$

and

$$V_{BA}^{tr} = k_{tr} (\cos\theta - \cos\theta_0)^2 \text{ (trigonometric potential)}$$

It can be seen that at $\theta \rightarrow \theta_0$ $V_{BA}^{tr} = k_{tr} (\theta - \theta_0)^2 \sin^2 \theta_0$, i.e., both versions of bond-angle potential coincide. CHARMM22 uses harmonic version of V_{BA} (Fig. 2), although the trigonometric one is computationally more efficient, because it does not require computation of \arccos functions or their partial derivatives.

bond angle		k_θ , kcal/(mol·deg ²)	θ_0 , deg	comments
(atom types)				
NH1	CT1 C	50.000	107.0000	! N-C α -C angle in backbone
HT	OT HT	55.000	104.5200	! H-O-H angle in water

Fig. 2 Parameterization of bond-angle potential (Eq. (2)) in CHARMM22 force field.

For water CHARMM22 offers the option to use instead Eq. (2) an Urey-Bradley (UB) potential to define the correct H-O-H bond angle. This potential has the same functional form as harmonic bond length potential with the parameter a being equal to equilibrium distance between water hydrogens. The UB potential is computationally less expensive and better reproduces vibrational frequencies of water.

III. Dihedral angle potential

Because of high energy constants in bond-length and bond-angle potentials, the associated degrees of freedom are effectively frozen at room temperature. In contrast, most other potentials have the energy constants comparable with the energy of thermal motion $k_B T$ and, as a result, they determine the structural transitions in proteins. The first of these “unfrozen” potentials is the dihedral potential, which plays a crucial role in local structure of proteins. The need for this potential arises because bond-length or bond-angle potentials as well as non-bonded potentials cannot easily describe the energetics of hydrocarbon molecules. For example, in four carbon butane C_4H_{10} dihedral angle potential is used to assign higher energy to *cis* conformation and describe properly the corresponding energy variation due to C-C bond rotation. The origin of dihedral potential is not completely understood, but repulsive interactions between overlapping bond orbitals and steric clashes between atoms (such as C_1 and C_4 in butane) appear to be contributing factors. The common functional form for dihedral angle potential V_{DA} is

$$V_{DA} = \frac{V_n}{2} [1 + \cos(n\phi - \phi_0)],$$

where V_n is the barrier height and n is the integer, which determines the periodicity of the potential (number of minima/maxima) in the interval $[0, 2\pi[$. The protein part of CHARMM22 uses $n=1, 2, 3$ or 4. Most common values are $n=1$ or 2.

Dihedral angle				V_n ,	n	ϕ_0	comments
(atom types)				kcal/mol			
C	CT1	NH1	C	0.2000	1	180.00	! backbone phi
NH1	C	CT1	NH1	0.6000	1	0.00	! backbone ksi
CT1	C	NH1	CT1	1.6000	1	0.00	! backbone omega
CA	CA	CA	CA	3.1000	2	180.00	! Phe side chain
H	OH1	CT2	CT1	0.4200	3	0.00	! Ser side chain

Fig. 3 Parameterization of dihedral angle potential in CHARMM22 force field.

The first three dihedral angle potentials in Fig.3 ($n=1$) have a single minimum at $\phi=0$ (*cis*) or $\phi=180^\circ$ (*trans*). The next dihedral angle potential ($n=2$) have two minima (at 0° and 180° or at 90° and 270° , if $\phi_0=0^\circ$). The potential with $n=3$ has three minima at 60° , 180° , 300° . In all cases, the minima are of the same depth. Combination of several dihedral potentials with $n=2$ and $n=3$ may be used to construct the potentials with uneven depths. These potentials are not used in proteins, but arise for nucleic acids. From Fig.3 it follows that the dihedral barriers are relatively low, except for ω backbone dihedral angle or that used for aromatic side chains. The parameters for V_{DA} are typically obtained from *ab initio* quantum calculations with additional optimization based on available experimental data.

IV. Improper dihedral angle potential

Improper dihedral angles are used to select the correct geometry or chirality of atoms. Consider four atoms i,j,k,l , among which j is linked covalently to i,l,k . The improper angle is defined as the angle between the (jl) line and the plane (ijk) (Fig. 4).

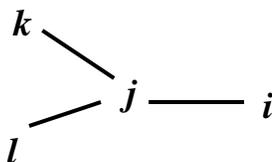


Fig. 4 Definition of improper angle for four atoms i,j,k,l .

The usual functional form of improper angle potential V_{imp} is

$$V_{imp} = k_{imp} (\psi - \psi_0)^2,$$

where k_{imp} determines the “stiffness” of the potential and ψ_0 is the equilibrium value. The example of improper angle parameterization is shown in Fig. 5.

Improper angle (atom types)	k_{imp} kcal / (mol · deg ²)	ψ_0	comments
O X X C	120.0000	0.0000	! l-k-i-j improper angle

Fig.5 Example of parameterization of improper angle potential in CHARMM22.

Fig. 5 can illustrate the use of improper angle to set the position of carbonyl oxygen O with respect to the plane C_α -C-N (in this case, X gives the atom types CT1 and NH1). In principle, this can be achieved with the usual dihedral angle potentials, but because the

planarity of C,O,C α ,N atoms must be strictly enforced, it is computationally cheaper to use stiff harmonic improper potential (note the high value of k_{imp}).

V. Non-bonded potential

Non-bonded potential includes van-der-Waals and electrostatic potentials. The van-der-Waals potentials take into account repulsion between atoms at small separations and weak attraction at larger distances. The common form of this potential for a pair of atoms i and j is given by a Lennard-Jones function V_{LJ} as

$$V_{LJ} = \frac{B_{ij}}{r_{ij}^{12}} - \frac{A_{ij}}{r_{ij}^6}$$

where B_{ij} and A_{ij} are the coefficients, which determine the depth and the location of energy minimum. The example of the parameters of van-der-Waals interactions is given in Fig. 5 for aliphatic carbons. The rules relating B_{ij} and A_{ij} coefficients to the parameters ε_i and σ_i in Fig. 5 are as follows

$$\begin{aligned} A_{ij} &= 2\varepsilon_{ij}\sigma_{ij}^6 \\ B_{ij} &= \varepsilon_{ij}\sigma_{ij}^{12} \\ \varepsilon_{ij} &= \sqrt{\varepsilon_i\varepsilon_j} \\ \sigma_{ij} &= \frac{1}{2}(\sigma_i + \sigma_j) \end{aligned}$$

atom type	ε_i kcal/mol	$\sigma_i/2$ Å				
CT1	0.000000	-0.020000	2.275000	0.000000	-0.010000	1.900000
CT2	0.000000	-0.055000	2.175000	0.000000	-0.010000	1.900000
CT3	0.000000	-0.080000	2.060000	0.000000	-0.010000	1.900000

Fig. 5 The parameters of van-der-Waals interactions for three aliphatic carbons.

The repulsion at small separations between atoms is associated with the Pauli exclusion principle, while weak attraction at larger distances is due to London dispersion interactions. The parameters ε_i and σ_i are usually obtained from X-ray diffraction experiments and from simulations of simple organic molecules, respectively. The important feature of V_{LJ} is its fast decay as $r \rightarrow \infty$ (due to r^6 term). For this reason van-der-Waals interactions are considered short-ranged and although the sum of van-der-Waals energies scales as N^2 , there are efficient schemes that reduce computational burden due to short-range character of this potential.

Electrostatic or Coulomb potential describes the interactions between pairs of partial charges. Its functional form is

$$V_{EL} = \frac{q_i q_j}{\epsilon(r_{ij}) r_{ij}},$$

where q_i and q_j are the partial charges on the atoms i and j and $\epsilon(r_{ij})$ is a distance dependent dielectric function. If solvent is treated explicitly in the simulations, the dielectric properties of the medium is taken into account automatically by explicitly computing all electrostatic interactions, therefore, $\epsilon(r_{ij})=1$. The need for r -dependent ϵ arises for implicit solvent models. The effective screening of electrostatic interactions due to water (the dielectric constant of water is 80) requires that V_{EL} decays faster than $1/r$. In the simplest case one can assume that $\epsilon(r)=r$ or consider the potential $\epsilon(r)=D \exp(kr)$, where k is the screening distance.

Because V_{EL} decays as r^{-1} , the electrostatic interactions are considered as long-ranged. For this reason it is difficult to devise methods reducing their N^2 scaling. One such method, Ewald sums, will be discussed in the future lectures. To quickly calculate V_{EL} in vacuum the following formula may be used

$$V_{EL} = \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r_{ij}} = 332 \text{ kcal } \dot{\text{A}} \text{ mol}^{-1} \frac{\tilde{q}_i \tilde{q}_j}{\tilde{r}_{ij}},$$

where \tilde{q}_i and \tilde{q}_j are expressed in the units of electron charge, r_{ij} is in Å, and ϵ_0 is the dielectric permittivity of vacuum. This formula indicates that the electrostatic interaction of two charges of 1e separated by the distance of 7 Å is about 50 kcal/mol.

VI. Different force fields

In these lectures we mainly considered CHARMM22 force field. However, many force fields are currently available. The most widely used force fields are

1. *AMBER (Assisted Model Building with Energy Refinement, amber.scripps.edu)*
2. *OPLS, OPLS-AA (Optimized Potentials for Liquid Simulations, zarbi.chem.yale.edu)*
3. *CHARMM (Chemistry at HARvard Macromolecular Mechanics, www.pharmacy.umaryland.edu/faculty/amackere/force_fields.htm)*
4. *GROMOS (GRONingen Molecular Simulation, www.igc.ethz.ch/gromos)*

It is important to keep in mind that these force fields are developed as whole and it is not possible to substitute or mix parts of different force fields.

The difficult question is related to the comparison of the performance of various force fields. In recent paper Brooks and coworkers compared the results of molecular dynamics

simulations of three large proteins using three force fields, CHARMM, OPLS-AA, and AMBER (*Journal of Computational Chemistry* **23**, 1045 (2002)). The simulations of 2 ns produced similar results. For example, the differences in the average accessible surface area, root-mean-squared deviation from the native state, secondary structure obtained from using different force fields were comparable with the normal variations between individual trajectories using a given force field.

Hummer and coworkers compared the conformational properties of polypeptides using CHARMM and AMBER force fields in the long (~10 ns) explicit water simulations (*JACS* **124**, 6563 (2002)). They found that, although the loop closure frequencies were similar, the distribution of structures differs. For example, the distribution of end-to-end distances was skewed towards larger distances in CHARMM. These studies suggest that open unfolded conformations of polypeptides show the evidence of force field dependence, but compact native-like structures have the properties largely independent on the particular force field. The likely explanation is that relatively shallow barriers separating structural states in the unfolded state depend on the details of specific parameterization, while self-averaging in the folded structures effectively cancels out the force field discrepancies.