

Contents

1	Force Field	3
1.1	Background	3
1.2	Covalent bonds	4
1.3	Angle potential	4
1.4	Van-der-Waals interactions	5
1.5	Switching function	6
1.5.1	Implicit evaluation of the switching function	7
1.5.2	Explicit evaluation of the switching function	7
1.6	Implicit solvent models	8
1.7	SASA solvation model	8
1.7.1	Parameters	9
1.7.2	Derivation of atomic force for SASA solvation model	9
1.7.3	Computational algorithm	12
1.8	Generalized Born Model	13
1.8.1	Original “Still” model	13
1.8.2	Hawkins-Cramer-Truhlar (HCT) model	14
1.8.3	Onufriev-Bashford-Case (OBC) model	15
1.8.4	ACE approximation for solvent accessible surface area	16
1.8.5	Parameters	16
1.8.6	Derivation of atomic forces	18
1.8.7	Computational algorithm	26
1.9	Long-range electrostatics	26
1.9.1	Ewald sumation	26

Chapter 1

Force Field

1.1 Background

The relationship between coordinates of atoms in a system and its energy is an essential part of any computational study based on atomic models. The potential energy function (force field) [1, 2] is used to calculate the potential energy of the system and to evaluate all the forces acting on all the atoms in the system. The force-field is the potential energy (i.e. function of coordinates) specified in terms of some constant parameters (parameters of the force field). A full-atomic potential energy function V is usually given by the sum of the bonded terms (V_b) and non-bonded terms (V_{nb}) [3–7], i.e.

$$V = V_b + V_{nb}, \quad (1.1)$$

where the bonded potential includes harmonic (covalent) bond part, harmonic angle and two types of torsion (dihedral) angles (proper and improper angles):

$$V_b = \sum_{bonds} \frac{k_b}{2} (b - b_0)^2 + \sum_{angles} \frac{k_\theta}{2} (\theta - \theta_0)^2 + \sum_{dihedrals} k_\phi (1 - \cos(n\phi - \phi_0)) + \sum_{impropers} \frac{k_\psi}{2} (\psi - \psi_0)^2 \quad (1.2)$$

For example, b and θ represent distance between two atoms and angle between two adjacent bonds; ϕ and ψ are dihedral (torsion) angles. These can be evaluated for all the atoms from their current positions. Also, k_b , k_θ , k_ϕ , and k_ψ are the spring constants, associated with bond vibrations, bending of bond angles, and conformational fluctuations in dihedral and improper angles around some equilibrium values b_0 , θ_0 , ϕ_0 , and ψ_0 , respectively. Importantly, each force-field has its own set of these parameters, which vary for different types of atoms [3, 8]. Some force-fields might also include Urey-Bradley angle corrections [3] and/or dihedral CMAP correction terms [4, 9, 10]. The non-bonded part of the potential energy function is represented by the electrostatic and Van-der-Waals potentials, i.e.

$$V_{nb} = \sum_{i,j} \left(\frac{q_i q_j}{4\pi\epsilon_0\epsilon r_{ij}} + \epsilon_{ij} \left[\left(\frac{R_{ij}^{min}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{ij}^{min}}{r_{ij}} \right)^6 \right] \right) \quad (1.3)$$

where r_{ij} is a distance between two interacting atoms, q_i and q_j are their charges; ϵ and ϵ_0 are electric and dielectric constant; $\epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j}$ and $R_{ij}^{min} = (R_i^{min} + R_j^{min})/2$ are Van-der-Waals parameters for atoms i and j . Early potential energy functions also included explicitly the hydrogen bonds potentials [3], which later have been dropped due to improved parametrization of electrostatics [4].

1.2 Covalent bonds

Covalent bonds are usually described by a harmonic potential:

$$V_{bonds} = \sum_{bonds} k_b(b - b_0)^2 \quad (1.4)$$

$$\nabla_l r_{ij} = \begin{cases} \frac{\vec{r}_{ij}}{r_{ij}} & \text{when } i = l \\ \frac{\vec{r}_{ji}}{r_{ij}} & \text{when } j = l \\ 0 & \text{when } i \neq l \text{ and } j \neq l \end{cases} \quad (1.5)$$

1.3 Angle potential

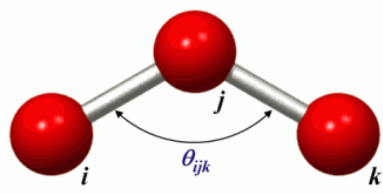


Figure 1.1: A schematic representation of a three atoms, interacting through angle potential. A triplet of atoms (i, j, k) and the current angle θ_{ijk} are shown.

As it is shown in Eq. 1.3, angle potential is described by following relation:

$$V_{angles} = \sum_{angles} \frac{k_\theta}{2} (\theta - \theta_0)^2 \quad (1.6)$$

Introducing a triplet (i, j, k) of atoms, involved in a particular angle (Fig. 1.1) and a set of angles in a system, A , one can re-write the equation above as:

$$V_{angles} = \sum_{(i,j,k) \in A} \frac{k_{ijk}^\theta}{2} (\theta_{ijk} - \theta_{ijk}^0)^2 \quad (1.7)$$

The force, acting on a particular atom l due to the angle potential is then:

$$\vec{f}_l = -\nabla_l V_{angles} = -\nabla_l \left(\sum_{(i,j,k) \in A} \frac{k_{ijk}^\theta}{2} (\theta_{ijk} - \theta_{ijk}^0)^2 \right) = - \sum_{(i,j,k) \in A} k_{ijk}^\theta (\theta_{ijk} - \theta_{ijk}^0) \nabla_l \theta_{ijk} \quad (1.8)$$

Since numerically it is easier to compute $\cos \theta_{ijk}$ and $\sin \theta_{ijk}$, let us switch to a cos and sin representation as follows:

$$\theta_{ijk} = \arccos(\cos \theta_{ijk}), \quad (1.9)$$

and

$$\nabla_l \theta_{ijk} = \nabla_l (\arccos(\cos \theta_{ijk})) = -\frac{1}{\sqrt{1 - \cos^2 \theta_{ijk}}} \nabla_l \cos \theta_{ijk} = -\frac{\nabla_l \cos \theta_{ijk}}{\sin \theta_{ijk}}. \quad (1.10)$$

Let us denote vectors that connect particles i , j and k in a triplet as $\vec{r}_{ji} = \vec{r}_i - \vec{r}_j$ and $\vec{r}_{jk} = \vec{r}_k - \vec{r}_j$. Then:

$$\cos \theta_{ijk} = \frac{\vec{r}_{ji} \cdot \vec{r}_{jk}}{|\vec{r}_{ji}| |\vec{r}_{jk}|} = \frac{\vec{r}_{ji} \cdot \vec{r}_{jk}}{r_{ji} r_{jk}} \quad (1.11)$$

Here, $\vec{r}_{ji} \cdot \vec{r}_{jk}$ is a dot product of vectors \vec{r}_{ji} and \vec{r}_{jk} . It is clear, that $\nabla_l \cos \theta_{ijk}$ will vanish if l does not belong to the triplet (i, j, k) . Considering all three cases ($l = i$, $l = j$ and $j = k$) separately and using Eq. 1.11 one gets:

$$\begin{aligned} \nabla_l \cos \theta_{ijk} \Big|_{l=i} &= \nabla_l \left(\frac{\vec{r}_{ji} \cdot \vec{r}_{jk}}{r_{ji} r_{jk}} \right) \Big|_{l=i} = \left(\frac{1}{r_{ji} r_{jk}} \nabla_l (\vec{r}_{ji} \cdot \vec{r}_{jk}) - \frac{\vec{r}_{ji} \cdot \vec{r}_{jk}}{r_{ji}^2 r_{jk}} \nabla_l r_{ji} \right) \Big|_{l=i} = \\ &= \frac{1}{r_{ji} r_{jk}} \vec{r}_{jk} - \frac{\vec{r}_{ji} \cdot \vec{r}_{jk}}{r_{ji}^2 r_{jk}} \frac{\vec{r}_{ji}}{r_{ji}} = \frac{1}{r_{ji}} \left[\frac{\vec{r}_{jk}}{r_{jk}} - \cos \theta_{ijk} \frac{\vec{r}_{ji}}{r_{ji}} \right], \\ \nabla_l \cos \theta_{ijk} \Big|_{l=k} &= \nabla_l \left(\frac{\vec{r}_{ji} \cdot \vec{r}_{jk}}{r_{ji} r_{jk}} \right) \Big|_{l=k} = \left(\frac{1}{r_{ji} r_{jk}} \nabla_l (\vec{r}_{ji} \cdot \vec{r}_{jk}) - \frac{\vec{r}_{ji} \cdot \vec{r}_{jk}}{r_{ji} r_{jk}^2} \nabla_l r_{jk} \right) \Big|_{l=k} = \\ &= \frac{1}{r_{ji} r_{jk}} \vec{r}_{ji} - \frac{\vec{r}_{ji} \cdot \vec{r}_{jk}}{r_{ji} r_{jk}^2} \frac{\vec{r}_{jk}}{r_{jk}} = \frac{1}{r_{jk}} \left[\frac{\vec{r}_{ji}}{r_{ji}} - \cos \theta_{ijk} \frac{\vec{r}_{jk}}{r_{jk}} \right], \text{ and} \\ \nabla_l \cos \theta_{ijk} \Big|_{l=j} &= \nabla_l \left(\frac{\vec{r}_{ji} \cdot \vec{r}_{jk}}{r_{ji} r_{jk}} \right) \Big|_{l=j} = \left(\frac{1}{r_{ji} r_{jk}} \nabla_l (\vec{r}_{ji} \cdot \vec{r}_{jk}) - \frac{\vec{r}_{ji} \cdot \vec{r}_{jk}}{r_{ji}^2 r_{jk}} \nabla_l r_{ji} - \frac{\vec{r}_{ji} \cdot \vec{r}_{jk}}{r_{ji} r_{jk}^2} \nabla_l r_{jk} \right) \Big|_{l=j} = \\ &= -\frac{1}{r_{ji} r_{jk}} \vec{r}_{jk} - \frac{1}{r_{ji} r_{jk}} \vec{r}_{ji} + \frac{\vec{r}_{ji} \cdot \vec{r}_{jk}}{r_{ji}^2 r_{jk}} \frac{\vec{r}_{ji}}{r_{ji}} + \frac{\vec{r}_{ji} \cdot \vec{r}_{jk}}{r_{ji} r_{jk}^2} \frac{\vec{r}_{jk}}{r_{jk}} = \\ &= \frac{1}{r_{ji}} \left[\cos \theta_{ijk} \frac{\vec{r}_{ji}}{r_{ji}} - \frac{\vec{r}_{jk}}{r_{jk}} \right] + \frac{1}{r_{jk}} \left[\cos \theta_{ijk} \frac{\vec{r}_{jk}}{r_{jk}} - \frac{\vec{r}_{ji}}{r_{ji}} \right] = \\ &= -\nabla_l \cos \theta_{ijk} \Big|_{l=i} - \nabla_l \cos \theta_{ijk} \Big|_{l=k} \end{aligned} \quad (1.12)$$

Summirizing Eqs. 1.8, 1.10 and 1.12, one can compute three components of force \vec{f}_i , \vec{f}_j and \vec{f}_k , acting on each atom in the angle triplet (i, j, k) due to the angle potential between atoms i , j and k using the following relations:

$$\begin{aligned} \vec{f}_i &= k_{ijk}^{\theta} (\theta_{ijk} - \theta_{ijk}^0) \left(-\frac{1}{\sin \theta_{ijk}} \right) \frac{1}{r_{ji}} \left[\cos \theta_{ijk} \frac{\vec{r}_{ji}}{r_{ji}} - \frac{\vec{r}_{jk}}{r_{jk}} \right] \\ \vec{f}_k &= k_{ijk}^{\theta} (\theta_{ijk} - \theta_{ijk}^0) \left(-\frac{1}{\sin \theta_{ijk}} \right) \frac{1}{r_{jk}} \left[\cos \theta_{ijk} \frac{\vec{r}_{jk}}{r_{jk}} - \frac{\vec{r}_{ji}}{r_{ji}} \right] \\ \vec{f}_j &= -\vec{f}_i - \vec{f}_k \end{aligned} \quad (1.13)$$

1.4 Van-der-Waals interactions

The formal definition of the Lennard-Jones potential that describes Van-der-Waals interactions reads:

$$V_{LJ} = \sum_{ij} V_{LJ}^{ij} = \sum_{ij} \varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right]. \quad (1.14)$$

Then, the force acting on the particle due to this potential is:

$$\vec{f}_{ij}^{LJ} = -\nabla_i V_{LJ}^{ij} = \varepsilon_{ij} \left[12 \frac{\sigma_{ij}^{12}}{r_{ij}^{13}} - 12 \frac{\sigma_{ij}^6}{r_{ij}^7} \right] \frac{\vec{r}_{ij}}{r_{ij}} = 12\varepsilon_{ij} \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 - 1 \right] \frac{\vec{r}_{ij}}{r_{ij}^2}. \quad (1.15)$$

Evaluation of the \vec{f}_{ij}^{LJ} requires N^2 binary forces to be computed. If a big atomistic system is considered, this can be a prohibitive factor for a long-timescale simulations. Most common work-around of this problem is to use the fact that the potential vanishes rapidly with distance. This allows one to cut the potential function off at some distance and, instead of computing potential for all N^2 binary pairs, consider only those particles that are close to the particle in question. This requires two more features to be implemented. First, a neighbours list - a list of particles that are close to the particle in question has to be constructed. Second, since the potential is not zero at any given distance, but rather asymptotically goes to zero, some changes have to be done to the potential function so that the value of the potential energy and its derivative is zero at the introduced cut-off distance. Later can be done using shifted or switched potential.

1.5 Switching function

Consider an arbitrary binary potential $V(\mathbf{r})$ ($\mathbf{r} = (\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N)$) that undergo a switching function $sw(\mathbf{r})$:

$$V^{sw}(\mathbf{r}) = V(\mathbf{r})sw(\mathbf{r}) = \sum_{i=1}^N \sum_{j=1}^N V_{ij}(r_{ij})sw_{ij}(r_{ij}) \quad (1.16)$$

According to the product rule, the atomic force acting on l -th particle due to this potential will then be:

$$\begin{aligned} \vec{f}_l(\mathbf{r}) &= -\nabla_l V^{sw} = -\nabla_l \left(\sum_{i=1}^N \sum_{j=1}^N V_{ij}(r_{ij})sw_{ij}(r_{ij}) \right) = \\ &= -\sum_{i=1}^N \sum_{j=1}^N (\nabla_l V_{ij}(r_{ij})) sw_{ij}(r_{ij}) + V_{ij}(r_{ij}) (\nabla_l sw_{ij}(r_{ij})) \end{aligned} \quad (1.17)$$

Computation of the atomic force due to the switched potential then can be done either explicitly or implicitly. In explicit approach, equation 1.17 should be expanded using the actual form of the $V(r)$. In implicit approach, all four terms of a product rule formula (Eq. 1.17) should be computed independently and then used to evaluate atomic force \vec{f}_l . The choice of the approach should be made considering computational performance, depending on the mathematical form of the potential $V(r)$.

The actual form of the switching function $sw(r)$ is given by [3]:

$$sw_{ij}(r_{ij}) = \begin{cases} 1 & \text{when } r_{ij} \leq r_{on} \\ \frac{(r_{off}^2 - r_{ij}^2)^2 (r_{off}^2 + 2r_{ij}^2 - 3r_{on}^2)}{(r_{off}^2 - r_{on}^2)^3} & \text{when } r_{on} < r_{ij} \leq r_{off} \\ 0 & \text{when } r_{ij} > r_{off} \end{cases} \quad (1.18)$$

where r_{on} and r_{off} are switching and cut-off distances respectively.

Since cases when $r_{ij} \leq r_{on}$ and when $r_{ij} > r_{off}$ are trivial, let us consider conditions when $r_{on} < r_{ij} \leq r_{off}$.

1.5.1 Implicit evaluation of the switching function

According to the product rule, one need to evaluate $V_{ij}(r_{ij})$, $\nabla_l V_{ij}(r_{ij})$, $sw_{ij}(r_{ij})$ and $\nabla_l sw_{ij}(r_{ij})$. Since all four quantities can be computed independently, here we will only consider evaluation of $sw_{ij}(r_{ij})$ and $\nabla_l sw_{ij}(r_{ij})$. Rearranging Eq. 1.18 ($r_{on} < r_{ij} \leq r_{off}$), we get:

$$sw_{ij}(r_{ij}) = \frac{2}{(r_{off}^2 - r_{on}^2)^3} (r_{off}^2 - r_{ij}^2)^2 \left(\frac{r_{off}^2 - 3r_{on}^2}{2} + r_{ij}^2 \right) \quad (1.19)$$

Since r_{on} and r_{off} are constant, one can use pre-computed values for:

$$C_1^{sw} = \frac{2}{(r_{off}^2 - r_{on}^2)^3}, C_2^{sw} = r_{off}^2, \text{ and } C_3^{sw} = \frac{r_{off}^2 - 3r_{on}^2}{2} \quad (1.20)$$

Then, Eq. 1.18 when $r_{on} < r_{ij} \leq r_{off}$ becomes:

$$sw_{ij}(r_{ij}) = C_1^{sw} (C_2^{sw} - r_{ij}^2)^2 (C_3^{sw} + r_{ij}^2) \quad (1.21)$$

Using Eq. 1.18 ($r_{on} < r_{ij} \leq r_{off}$), we also can compute $\nabla_l sw_{ij}(r_{ij})$:

$$\begin{aligned} \nabla_l sw_{ij}(r_{ij}) &= \frac{1}{(r_{off}^2 - r_{on}^2)^3} \left[2 (r_{off}^2 - r_{ij}^2) (-2r_{ij}) (r_{off}^2 + 2r_{ij}^2 - 3r_{on}^2) + (r_{off}^2 - r_{ij}^2)^2 4r_{ij} \right] \frac{\vec{r}_{ij}}{r_{ij}} = \\ &= \frac{1}{(r_{off}^2 - r_{on}^2)^3} \left[-4r_{off}^4 - 8r_{off}^2 r_{ij}^2 + 12r_{off}^2 r_{on}^2 + 4r_{ij}^2 r_{off}^2 + 8r_{ij}^4 - 12r_{ij}^2 r_{on}^2 + 4r_{off}^4 - 8r_{off}^2 r_{ij}^2 + 4r_{ij}^4 \right] \frac{\vec{r}_{ij}}{r_{ij}} = \\ &= \frac{1}{(r_{off}^2 - r_{on}^2)^3} \left[12r_{ij}^4 - 12r_{ij}^2 r_{off}^2 - 12r_{ij}^2 r_{on}^2 + 12r_{off}^2 r_{on}^2 \right] \frac{\vec{r}_{ij}}{r_{ij}} = \\ &= \frac{12}{(r_{off}^2 - r_{on}^2)^3} (r_{ij}^2 - r_{off}^2) (r_{ij}^2 - r_{on}^2) \frac{\vec{r}_{ij}}{r_{ij}} \end{aligned} \quad (1.22)$$

Using pre-computed values for:

$$C^{dsw} = \frac{12}{(r_{off}^2 - r_{on}^2)^3}, r_{off}^2, \text{ and } r_{on}^2 \quad (1.23)$$

Equation 1.22 becomes:

$$\nabla_l sw_{ij}(r_{ij}) = C^{dsw} (r_{ij}^2 - r_{off}^2) (r_{ij}^2 - r_{on}^2) \frac{\vec{r}_{ij}}{r_{ij}} \quad (1.24)$$

1.5.2 Explicit evaluation of the switching function

Let us expand Eq. 1.16 for the Lennard-Jones potential.

1.6 Implicit solvent models

Accurate representation of the solvent environment is important for realistic biomolecular modeling. The most straightforward approach is to expose the biomolecule to explicitly defined water molecules. However, to avoid boundary effects, the amount of water molecules added to the system is very large. Typically, in all-atom MD simulations in explicit water, 90% of all the degrees of freedom correspond to the water environment. This results in the increased system size and the computational complexity, significantly slowing down the computational procedures, thus, reducing reachable timescales of simulations. This might prohibit the exploration of the biologically relevant transitions in biomolecules. Alternatively, interactions of protein and water molecules can be described by mechanical “kicks” of the protein atoms and Coulomb interactions that include electrostatic “screening” effect and entropy-driven hydrophobic effect. This approach, called implicit solvent representation [4, 11], involves introducing Langevin equations of motion, where random forces represent stochastic water kicks experienced by the protein molecule, and additional potential energy function to model electrostatic solvation effects. Implicit solvation models have several advantages over explicit water representations including lower computational cost, better thermodynamic sampling and more straightforward methodology for the free energy estimations.

In most implicit solvent models, the total potential energy of the solvated molecule is given by

$$V_{tot} = V_{vac} + \Delta G_{solv} \quad (1.25)$$

where V_{vac} is the potential energy of the molecule in vacuum (i.e. in the absence of water environment) and ΔG_{solv} is defined as the free energy of transferring the molecule from vacuum to the solvent (solvation free energy). Typically, V_{vac} is computed using the molecular force field (Eqs. 1.1-1.3), whereas ΔG_{solv} depends on the particular implicit solvent model used. ΔG_{solv} can be decomposed into the polar and non-polar parts, which represent the hydrophobic effect and electrostatic effect of the solvent, respectively. There are several different approaches for implicit water representation [4, 11] including Solvent Accessible Surface Area (SASA) model [12, 13] and Generalized Born (GB) model, which is based on the Generalized Born approximation of the Poisson-Boltzmann equation [14–19]. Some other methods [20] combine different descriptions of implicit solvation into hybrid models [21].

1.7 SASA solvation model

In Solvent Accessible Surface Area solvation energy approximation, it is assumed that the mean solvation energy $\Delta G_{solv}(\mathbf{r})$ ($\mathbf{r} = (\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N)$) is linearly related to the solvent accessible surface area of a solute [12, 13, 22]:

$$\Delta G_{solv}^{SASA}(\mathbf{r}) = \sum_{i=1}^N \sigma_i A_i(\mathbf{r}), \quad (1.26)$$

where σ_i is an atomic solvation parameter and $A_i(\mathbf{r})$ is a SASA of atom i . Later is computed using approximate analytical expression:

$$A_i(\mathbf{r}) = S_i \prod_{\substack{j=1 \\ j \neq i}}^N \left[1 - \frac{p_i p_{ij} b_{ij}(r_{ij})}{S_i} \right] \quad (1.27)$$

Table 1.1: CHARMM Atom Types and Mean Solvation Parameters (data are taken from [12]).

Atom Type	R_{vdW}^{min} , Å	R_i , Å	p_i	σ_i , kcal/molÅ ²	Description
C	2.1	1.72	1.554	0.012	Carbonyl carbon
CH1E	2.365	1.80	1.276	0.012	Extended aliphatic carbon with 1 hydrogen
CH2E	2.235	1.90	1.045	0.012	Extended aliphatic carbon with 2 hydrogens
CH3E	2.165	2.00	0.880	0.012	Extended aliphatic carbon with 3 hydrogens
CR1E	2.1	1.80	1.073	0.012	Extended aromatic carbon with 1 hydrogen
CR	2.1	1.80	1.073	0.012	Extended aromatic carbon with 1 hydrogen
NH1	1.6	1.55	1.028	-0.060	Amide nitrogen
NR	1.6	1.55	1.028	-0.060	Aromatic nitrogen with no hydrogens
NH2	1.6	1.60	1.215	-0.060	Nitrogen with two hydrogens
NH3	1.6	1.60	1.215	-0.060	Nitrogen with three hydrogens
NC2	1.6	1.55	1.028	-0.060	Guanidinium nitrogen
N	1.6	1.55	1.028	-0.060	Proline nitrogen
OH1	1.6	1.52	1.080	-0.060	Hydroxyl oxygen
O	1.6	1.50	0.926	-0.060	Carbonyl oxygen
OC	1.6	1.70	0.922	-0.060	Carboxyl oxygen
S	1.89	1.80	1.121	0.012	Sulphur
SH1E	1.89	1.80	1.121	0.012	Extended sulphur with 1 hydrogen
H	0.8	1.10	1.128	0.000	Polar hydrogen
HC	0.6	1.10	1.128	0.000	Polar hydrogen (in Arg, Lys and N-term)

where S_i is a SASA of an isolated atom i of radius R_i :

$$S_i = 4\pi (R_i + R_{probe})^2 \quad (1.28)$$

R_{probe} is a radius of a solvent probe. In Eq. 1.27, $b_{ij}(r_{ij})$ represents the SASA, removed from S_i due to the overlap between atoms i and j , r_{ij} is a distance between atoms i and j ($r_{ij} = |\vec{r}_i - \vec{r}_j|$). b_{ij} is given by the following equation:

$$b_{ij} = \begin{cases} 0 & \text{when } r_{ij} > R_i + R_j + 2R_{probe} \\ \pi (R_i + R_{probe}) (R_i + R_j + 2R_{probe} - r_{ij}) \left(1 + \frac{R_j - R_i}{r_{ij}}\right) & \text{otherwise} \end{cases} \quad (1.29)$$

1.7.1 Parameters

Atom parameters p_i , connectivity parameters p_{ij} have been selected to reproduce exact SASA with $R_{probe} = 1.4\text{\AA}$. The parameter p_{ij} is 0.8875 if the atoms i and j are covalently bonded and 0.3516 otherwise. The values for σ_i , R_i and p_i for different atom types are given in Table 1.1 [12].

1.7.2 Derivation of atomic force for SASA solvation model

Let us first assume, that only pairs of atoms (i, j) , for which condition $r_{ij} < R_i + R_j + 2R_{probe}$ is satisfied. In this case:

$$b_{ij} = \pi (R_i + R_{probe}) (R_i + R_j + 2R_{probe} - r_{ij}) \left(1 + \frac{R_j - R_i}{r_{ij}}\right) \quad (1.30)$$

and summation and product in Eqs. 1.26 and 1.27 are taken over pairs, for which $b_{ij}(r_{ij}) \neq 0$. To evaluate the gradient of a solvation energy ($\vec{f}_l(\mathbf{r}) = \nabla_l \Delta G_{solv}^{SASA}(\mathbf{r})$), we will need to find $\nabla_l b_{ij}(r_{ij})$.

Later are not zero only if $l = i$ or $l = j$. Denoting $R_{i,pr} = R_i + R_{probe}$, we will have:

$$b_{ij} = \pi R_{i,pr} (R_{i,pr} + R_{j,pr} - r_{ij}) \left(1 + \frac{R_{j,pr} - R_{i,pr}}{r_{ij}} \right) \quad (1.31)$$

Then,

$$\begin{aligned} \nabla_i b_{ij} &= \pi R_{i,pr} \nabla_i \left[(R_{i,pr} + R_{j,pr} - r_{ij}) \left(1 + \frac{R_{j,pr} - R_{i,pr}}{r_{ij}} \right) \right] = \\ &= \pi R_{i,pr} \left[-\frac{\vec{r}_{ij}}{r_{ij}^2} \left(1 + \frac{R_{j,pr} - R_{i,pr}}{r_{ij}} \right) + (R_{i,pr} + R_{j,pr} - r_{ij}) \frac{R_{j,pr} - R_{i,pr}}{-r_{ij}^2} \frac{\vec{r}_{ij}}{r_{ij}} \right] = \\ &= \pi R_{i,pr} \left[-1 - \frac{R_{j,pr} - R_{i,pr}}{r_{ij}} - (R_{i,pr} + R_{j,pr}) \frac{R_{j,pr} - R_{i,pr}}{-r_{ij}^2} + \frac{R_{j,pr} - R_{i,pr}}{r_{ij}} \right] \frac{\vec{r}_{ij}}{r_{ij}} \end{aligned} \quad (1.32)$$

Which follows:

$$\nabla_i b_{ij} = -\pi R_{i,pr} \left[1 + \frac{R_{j,pr}^2 - R_{i,pr}^2}{r_{ij}^2} \right] \frac{\vec{r}_{ij}}{r_{ij}^3} \quad (1.33)$$

Similarly:

$$\nabla_j b_{ij} = \pi R_{i,pr} \left[1 + \frac{R_{j,pr}^2 - R_{i,pr}^2}{r_{ij}^2} \right] \frac{\vec{r}_{ij}}{r_{ij}^3} \quad (1.34)$$

Let us denote:

$$B_{ij} = 1 - \frac{p_i p_{ij} b_{ij}(r_{ij})}{S_i} \quad (1.35)$$

then

$$\begin{aligned} \nabla_l B_{ij} &= -\frac{p_i p_{ij}}{S_i} \nabla_l b_{ij}(r_{ij}) \\ A_i(\mathbf{r}) &= S_i \prod_{j \neq i}^N B_{ij}(r_{ij}) \\ \Delta G_{solv}^{SASA}(\mathbf{r}) &= \sum_{i=1}^N \sigma_i S_i \prod_{\substack{j=1 \\ j \neq i}}^N B_{ij}(r_{ij}) \end{aligned} \quad (1.36)$$

Now, we can compute atomic force (i.e. gradient of the solvation energy):

$$\vec{f}_l(\mathbf{r}) = -\nabla_l \Delta G_{solv}^{SASA}(\mathbf{r}) = -\nabla_l \sum_{i=1}^N \sigma_i S_i \prod_{\substack{j=1 \\ j \neq i}}^N B_{ij}(r_{ij}) \quad (1.37)$$

Since $\nabla_l B_{ij}(r_{ij}) \neq 0$ only if $l = i$ or $l = j$:

$$\vec{f}_l(\mathbf{r}) = -\sigma_l S_l \nabla_l \prod_{\substack{j=1 \\ j \neq l}}^N B_{lj}(r_{lj}) - \sum_{\substack{i=1 \\ i \neq l}}^N \sigma_i S_i \nabla_l \prod_{\substack{j=1 \\ j \neq i}}^N B_{ij}(r_{ij}) = \vec{f}_1(\mathbf{r}) + \vec{f}_2(\mathbf{r}) \quad (1.38)$$

The first term ($\vec{f}_1(\mathbf{r})$) in Eq. 1.38 can be decomposed as follows:

$$\begin{aligned}
 \vec{f}_1(\mathbf{r}) &= -\sigma_l S_l \nabla_l \prod_{\substack{j=1 \\ j \neq l}}^N B_{lj}(r_{lj}) = -\sigma_l S_l \sum_{\substack{j=1 \\ j \neq l}}^N (\nabla_l B_{lj}(r_{lj})) \prod_{\substack{k=1 \\ k \neq l \\ k \neq j}}^N B_{lk}(r_{lk}) = \\
 &= -\sigma_l S_l \sum_{\substack{j=1 \\ j \neq l}}^N (\nabla_l B_{lj}(r_{lj})) \frac{B_{lj}(r_{lj})}{B_{lj}(r_{lj})} \prod_{\substack{k=1 \\ k \neq l \\ k \neq j}}^N B_{lk}(r_{lk}) = \\
 &= -\sigma_l S_l \sum_{\substack{j=1 \\ j \neq l}}^N (\nabla_l B_{lj}(r_{lj})) \frac{\prod_{\substack{k=1 \\ k \neq l}}^N B_{lk}(r_{lk})}{B_{lj}(r_{lj})}
 \end{aligned} \tag{1.39}$$

Changing index from j to i , we get:

$$\vec{f}_1(\mathbf{r}) = - \sum_{\substack{i=1 \\ i \neq l}}^N (\nabla_l B_{li}(r_{li})) \frac{V_l(\mathbf{r})}{B_{li}(r_{li})} \tag{1.40}$$

where

$$V_l(\mathbf{r}) = \sigma_l S_l \prod_{\substack{k=1 \\ k \neq l}}^N B_{lk}(r_{lk}) \tag{1.41}$$

Note, that:

$$\sum_{l=1}^N V_l(\mathbf{r}) = \Delta G_{solv}^{SASA}(\mathbf{r}) \tag{1.42}$$

Let us now consider the second term ($\vec{f}_2(\mathbf{r})$) in Eq. 1.38:

$$\vec{f}_2(\mathbf{r}) = - \sum_{\substack{i=1 \\ i \neq l}}^N \sigma_i S_i \nabla_l \prod_{\substack{j=1 \\ j \neq i}}^N B_{ij}(r_{ij}) \tag{1.43}$$

Since $j \neq i$, only one term inside the product will depend on \vec{r}_l , i.e. when $j = l$:

$$\begin{aligned}
 \vec{f}_2(\mathbf{r}) &= - \sum_{\substack{i=1 \\ i \neq l}}^N \sigma_i S_i (\nabla_l B_{il}(r_{il})) \prod_{\substack{j=1 \\ j \neq i \\ j \neq l}}^N B_{ij}(r_{ij}) = \\
 &= - \sum_{\substack{i=1 \\ i \neq l}}^N \sigma_i S_i (\nabla_l B_{il}(r_{il})) \frac{B_{il}(r_{il})}{B_{il}(r_{il})} \prod_{\substack{j=1 \\ j \neq i \\ j \neq l}}^N B_{ij}(r_{ij}) = \\
 &= - \sum_{\substack{i=1 \\ i \neq l}}^N \sigma_i S_i (\nabla_l B_{il}(r_{il})) \frac{\prod_{\substack{j=1 \\ j \neq i}}^N B_{ij}(r_{ij})}{B_{il}(r_{il})}
 \end{aligned} \tag{1.44}$$

Finally:

$$\vec{f}_2(\mathbf{r}) = - \sum_{\substack{i=1 \\ i \neq l}}^N (\nabla_l B_{il}(r_{il})) \frac{V_i(\mathbf{r})}{B_{il}(r_{il})} \quad (1.45)$$

Combining Eqs. 1.40 and 1.45, we get:

$$\vec{f}_l(\mathbf{r}) = - \sum_{\substack{i=1 \\ i \neq l}}^N \left[\frac{(\nabla_l B_{li}(r_{li})) V_l(\mathbf{r})}{B_{li}(r_{li})} + \frac{(\nabla_l B_{il}(r_{il})) V_i(\mathbf{r})}{B_{il}(r_{il})} \right] \quad (1.46)$$

1.7.3 Computational algorithm

The computational procedure for atomic forces can be decomposed into the following steps:

Step 1: Find all atom pairs, for which:

$$r_{ij} = |\vec{r}_i - \vec{r}_j| < R_i + R_j + 2R_{probe} = R_{i,pr} + R_{j,pr} \quad (1.47)$$

Step 2: For all pairs (i, j) found, compute following values:

$$\begin{aligned} B_{ij} &= 1 - \frac{p_i p_{ij}}{S_i} b_{ij}(r_{ij}) = 1 - \frac{p_i p_{ij}}{S_i} \pi R_{i,pr} (R_{i,pr} + R_{j,pr} - r_{ij}) \left(1 + \frac{R_{j,pr} - R_{i,pr}}{r_{ij}} \right) \\ B_{ji} &= 1 - \frac{p_j p_{ij}}{S_j} b_{ji}(r_{ij}) = 1 - \frac{p_j p_{ij}}{S_j} \pi R_{j,pr} (R_{i,pr} + R_{j,pr} - r_{ij}) \left(1 + \frac{R_{i,pr} - R_{j,pr}}{r_{ij}} \right) \\ \nabla_i B_{ij} &= -\frac{p_i p_{ij}}{S_i} \nabla_i b_{ij}(r_{ij}) = \frac{p_i p_{ij}}{S_i} \pi R_{i,pr} \left(1 + \frac{R_{j,pr}^2 - R_{i,pr}^2}{r_{ij}^2} \right) \frac{\vec{r}_{ij}}{r_{ij}} \\ \nabla_j B_{ji} &= -\frac{p_j p_{ij}}{S_j} \nabla_j b_{ji}(r_{ij}) = \frac{p_j p_{ij}}{S_j} \pi R_{j,pr} \left(1 + \frac{R_{i,pr}^2 - R_{j,pr}^2}{r_{ij}^2} \right) \frac{\vec{r}_{ij}}{r_{ij}} \end{aligned} \quad (1.48)$$

Note, that $B_{ij} \neq B_{ji}$ and $\nabla_i B_{ij} \neq \nabla_j B_{ji}$. However, $\nabla_j B_{ij} = -\nabla_i B_{ji}$ as it is follows from Eqs. 1.33 and 1.34.

Step 3: Using values for B_{ij} and Eq. 1.41, compute values of V_i for all atoms:

$$V_i(\mathbf{r}) = \sigma_i S_i \prod_{\substack{j=1 \\ j \neq i}}^N B_{ij}(r_{ij}) \quad (1.49)$$

At this point, total solvation energy ΔG_{solv}^{SASA} can be evaluated using Eq. 1.42:

$$\Delta G_{solv}^{SASA} = \sum_{i=1}^N V_i(\mathbf{r}) \quad (1.50)$$

Step 4: Using Eq. 1.46 and values, computed at steps 2 and 3, we can now evaluate the force $\vec{f}_i(\mathbf{r})$ acting on a particle i due to the solvation potential ΔG_{solv}^{SASA} :

$$\vec{f}_i(\mathbf{r}) = - \sum_{\substack{j=1 \\ j \neq i}}^N \left[\frac{(\nabla_i B_{ij}) V_i(\mathbf{r})}{B_{ij}} + \frac{(\nabla_j B_{ji}) V_j(\mathbf{r})}{B_{ji}} \right] \quad (1.51)$$

1.8 Generalized Born Model

In GB model, the solute molecule (protein) surrounded by the dielectric media (water) experiences the screening effect. Since water molecules are not described explicitly, plain Coulomb potential is not accurate at describing the electrostatic interactions (Eq. 1.3). In GB model, accurate evaluation of electrostatics can be achieved using the Poisson-Boltzmann model or its less computationally expensive Generalized Born approximation. The free energy of solvation in GB model is represented by the following correction for the electrostatic energy term,

$$\Delta G_{solv}^{GB} = -\frac{k_c}{2} \left(\frac{1}{\varepsilon_{in}} - \frac{1}{\varepsilon_{out}} \right) \sum_{i=1}^N \sum_{j=1}^N \frac{q_i q_j}{\sqrt{r_{ij}^2 + \alpha_i \alpha_j \exp \left[-\frac{r_{ij}^2}{\gamma \alpha_i \alpha_j} \right]}} \quad (1.52)$$

where $k_c = 1/4\pi\epsilon_0$ is the Coulomb constant, ε_{in} and ε_{out} are the dielectric constants for the solute and solvent respectively, q_i is the charge on the atom i , r_{ij} is a distance between i -th and j -th atoms, and γ is a numerical parameter (normally, $\gamma=4$). The Born radii of atoms, α_i , characterize the degree of burial of the particular atom into the solute. When Born radii of atoms are small, the term with exponent in the denominator in Eq. 1.52 can be neglected, and then sum of electrostatic term (Eq. 1.3) and Generalized Born term for this pair of atoms will be identical to the Coulomb potential with dielectric constant ε_{out} . For very large α 's, ΔG_{solv}^{GB} will be small and can be neglected.

The original GB model (Still model) [14, 15, 21] was based on the idea, that any atom displaces the volume of solute equal the atom's Van-der-Waals volume. Later, this approach was improved to account for the gaps between the Van-der-Waals volume of atoms that are too small to include water molecules (HTC model)[18]. Further improvement [19] was made for a better description of a macromolecules, where the electrostatic potential for a deeply buried atoms should not be corrected by ΔG_{solv}^{GB} (OBC model). Other variants of the Generalized Born model (GBSW/GBMV/GBMV2 models) are based on the numerical integration over the medium to find the degree of burial for a particular atom [16, 17, 23, 24].

1.8.1 Original "Still" model

The empirical expression first proposed by Still relates the "generalized Born" radius (α_i) of an atom in a specific molecular environment to the polarization energy of that atom in the same environment through the classical equation for the Born polarization free energy [14, 15, 21]:

$$\alpha_i = -\frac{k_c}{2} \frac{1}{G_{pol,i}} \quad (1.53)$$

The analytical formula for atomic polarization energy utilized by Still is based on the premise that the polarization of an external medium (such as water) due to a charge is reduced proportionally to the volume of displaced dielectric around the charge. This displaced dielectric originates from the volume occupied by atoms surrounding the point charge of interest. The reduction of polarization is inversely proportional to the distance (from the charge to the displaced volume elements) raised to the fourth power, as derived from the classical charge-induced dipole interaction [25]. It has also proven useful to separate the contributions to the burial of a specific point charge based on the chemical topology. In the model of Still, the relative contribution of surrounding atoms toward the reduction of polarization

energy is partitioned based on topology as follows: atoms bonded to the charge of interest, atoms within a bond angle with the charge of interest, and atoms related to the charge only through nonbonded interactions [15, 21]. A linearized form of Still's original empirical formula:

$$G_{pol,i} = \left(\frac{1}{\varepsilon_{in}} - \frac{1}{\varepsilon_{out}} \right) \left[\frac{1}{\lambda} \left(-\frac{k_c}{2R_i} \right) + P_1 \left(\frac{k_c}{2R_i^2} \right) + \sum_j^{bond} \frac{P_2 V_j}{r_{ij}^4} + \sum_j^{angle} \frac{P_3 V_j}{r_{ij}^4} + \sum_j^{nonbond} \frac{P_4 V_j}{r_{ij}^4} CCF \right], \quad (1.54)$$

where CCF is a Close Contact Function, defined as:

$$CCF = \begin{cases} 1.0 & \text{when } \left(\frac{r_{ij}}{R_i + R_j} \right)^2 > \frac{1}{P_5} \\ \left(\frac{1}{2} \left[1.0 - \cos \left(\left(\frac{r_{ij}}{R_i + R_j} \right)^2 P_5 \pi \right) \right] \right)^2 & \text{when } \left(\frac{r_{ij}}{R_i + R_j} \right)^2 \leq \frac{1}{P_5} \end{cases} \quad (1.55)$$

In this expression, r_{ij} is the distance from the point charge of interest, i , and an uncharged atom j in the molecule; R_i is the van der Waals radius of atom i (taken as the radius parameter R_{min} for atom i in the CHARMM parameters file); ε_{in} and ε_{out} are the dielectric constants of the solute and solvent respectively; and $V_j = 4/3\pi R_j^3$ is the atomic volume of atom j . The effective reduction in atomic volume due to neighboring atoms is absorbed to some extent by the fitting parameters. The fitting parameters are P_1 , P_2 , P_3 , P_4 , P_5 , and λ . From this expression for the solvation energy of a unit charge ion, the Born radius is computed using the Born equation (Eq. 1.54).

1.8.2 Hawkins-Cramer-Truhlar (HCT) model

Similarly to Still's approach [14], the generalized Born approximation, proposed by G. Hawkins, C. Cramer and D. Truhlar (hence the name 'HCT'), is based on the pairwise descreening approximation. The Born radii in this case can be written as:

$$\alpha_i^{-1} = \rho_i^{-1} - \sum_{\substack{j=1 \\ j \neq i}}^N \int_{\rho_i}^{\infty} \frac{dr}{r^2} \mathcal{H}_{ij}(r_{ij}, \rho_j) = \rho_i^{-1} - I_i, \quad (1.56)$$

where \mathcal{H}_{ij} is the fraction of the area of a sphere of radius r centered at atom j that is shielded by a sphere of scaled radius $S_{ij}\rho_j$ at a distance r_{ij} [18, 26]. Interestingly, the scale factor S_{ij} originally was introduced as dependent on both atom types of i -th and j -th atom [18]. Later, this approach was dropped for the sake of only j -th atom type dependence. Although this transition was never documented, the later approach is used in most implementations, including Tinker and Gromacs. Because of that, we will use the notation S_j from now on. The scaling factors and atom radii for most force-fields can be found, for instance, in Gromacs topologies. Also note, that in Eq. 1.56, the sum does not include the i -th atom since the self-screening is already taken into account (i.e. if the sum is zero, $\alpha_i = \rho_i$, which corresponds to the isolated atom). For the same reason, the lower limit in integral is ρ_i . It was shown that using slightly reduced values for $\rho_i = \rho_{0i} - \rho_{dd}$, instead of ρ_{0i} gives a better agreement with the corresponding Poisson-Boltzmann calculations. The parameter ρ_{dd} called dielectric displacement and usually set to $\rho_{dd} = 0.09\text{\AA}$. Note, that at least in the Gromacs topologies unadjusted radii (ρ_{0i}) are listed.

It is possible to calculate \mathcal{H}_{ij} analytically [18, 26, 27]. Suppose that sphere i has a radius r and that the sphere j has radius $S_j\rho_j$. If the two spheres intersect then $r - S_j\rho_j \leq r_{ij} \leq r + S_j\rho_j$, and the fraction of the surface area of the sphere i eclipsed by a sphere j can be calculated by:

$$\mathcal{H}_{ij} = \frac{1}{2}(1 - \cos \theta_j^{(i)}) = \frac{1}{2} - \frac{r_{ij}^2 + r^2 - (S_j\rho_j)^2}{4r_{ij}r}, \quad (1.57)$$

where $\theta_j^{(i)}$ is the angle between the line connecting the centers of the spheres and the line, connecting the center of the sphere i and one of the spheres intersection points. The expansion of the cosine of this angle in Eq. 1.57 comes from the cosine theorem, since $\theta_j^{(i)}$ is one of the angles in a triangle made by two spheres radii r and $S_j\rho_j$ and the line connecting the centers of the spheres r_{ij} . If the spheres don't intersect (i.e. when $r_{ij} > r + S_j\rho_j$), the eclipsed surface area should be zero.

Using the relation for \mathcal{H}_{ij} (Eq. 1.57), we can now found the Born radii (Eq. 1.56):

$$\begin{aligned} \alpha_i^{-1} &= \rho_i^{-1} - \frac{1}{2} \sum_{\substack{j=1 \\ j \neq i}}^N \int_{L_{ij}}^{U_{ij}} \left(\frac{1}{r^2} - \frac{r_{ij}}{2r^3} - \frac{1}{2r_{ij}r} + \frac{(S_j\rho_j)^2}{2r_{ij}r^3} \right) dr = \\ &= \rho_i^{-1} - \frac{1}{2} \sum_{\substack{j=1 \\ j \neq i}}^N \left[\frac{1}{L_{ij}} - \frac{1}{U_{ij}} + \frac{r_{ij}}{4} \left(\frac{1}{U_{ij}^2} - \frac{1}{L_{ij}^2} \right) + \frac{1}{2r_{ij}} \ln \frac{L_{ij}}{U_{ij}} + \frac{(S_j\rho_j)^2}{4r_{ij}} \left(\frac{1}{L_{ij}^2} - \frac{1}{U_{ij}^2} \right) \right] \end{aligned} \quad (1.58)$$

The lower and upper limits in integral (L_{ij} and U_{ij}) depend on whether or not spheres intersect. The relations for these values can be written as:

$$L_{ij} = \begin{cases} 1 & \text{if } r_{ij} + S_j\rho_j \leq \rho_i \\ \rho_i & \text{if } r_{ij} - S_j\rho_j \leq \rho_i < r_{ij} + S_j\rho_j \\ r_{ij} - S_j\rho_j & \text{if } \rho_i \leq r_{ij} - S_j\rho_j \end{cases} \quad (1.59)$$

and

$$U_{ij} = \begin{cases} 1 & \text{if } r_{ij} + S_j\rho_j \leq \rho_i \\ r_{ij} + S_j\rho_j & \text{if } \rho_i \leq r_{ij} - S_j\rho_j \end{cases} \quad (1.60)$$

Note, that the value of 1 in case spheres do not intersect ($r_{ij} + S_j\rho_j \leq \rho_i$) is arbitrary and serve the purpose to make the integral zero.

1.8.3 Onufriev-Bashford-Case (OBC) model

A. Onufriev, D. Bashford and A. Case claimed that the HTC model routinely underestimates the effective Born radii [19, 28]. To work around this issue, they proposed to rescale the Born radii α_i with the rescaling parameters being proportional to the degree of atom's burial, as quantified by the value of I_i in Eq. 1.56. The rescaling function to replace the original expression for the effective Born radii (Eq. 1.56) was chosen to be [28]:

$$\alpha_i^{-1} = \rho_i^{-1} - \rho_{0i}^{-1} \tanh(\alpha\Psi_i - \beta\Psi_i^2 + \gamma\Psi_i^3), \quad (1.61)$$

where $\rho_i = \rho_{0i} - \rho_{dd}$, ρ_{dd} is dielectric displacement (commonly used value is 0.09Å) [14, 28]; $\Psi_i = I_i\rho_i$; α , β , γ are treated as adjustable dimensionless parameters to be optimized [28]. Two sets of values

for these parameters were found in original paper: $\alpha = 0.8$, $\beta = 0$, $\gamma = 2.91$ and $\alpha = 1.0$, $\beta = 0.8$, $\gamma = 4.85$. The actual form of the integral I_i follows from the Eq. 1.58:

$$\Psi_i = \rho_i I_i = \rho_i \frac{1}{2} \sum_{\substack{j=1 \\ j \neq i}}^N \left[\frac{1}{L_{ij}} - \frac{1}{U_{ij}} + \frac{r_{ij}}{4} \left(\frac{1}{U_{ij}^2} - \frac{1}{L_{ij}^2} \right) + \frac{1}{2r_{ij}} \ln \frac{L_{ij}}{U_{ij}} + \frac{(S_j \rho_j)^2}{4r_{ij}} \left(\frac{1}{L_{ij}^2} - \frac{1}{U_{ij}^2} \right) \right], \quad (1.62)$$

with the values for lower and upper limits L_{ij} and U_{ij} identical to HCT model (Eqs. 1.59 and 1.60).

1.8.4 ACE approximation for solvent accessible surface area

The non-polar part of the potential in so-called GB/SA approach usually estimated using SASA approach. The formalism used can be the same as in the SASA implicit solvent but having different parametrization since, compared to the stand-alone SASA implicit solvent, it no longer required to take into account electrostatic effects of solvation. For this matter, single σ parameter is usually used for all atoms. This approach was originally proposed by Still [14, 21, 29]. Another way of estimating non-polar SASA part of the GB/SA potential based on the fact that the Born radii α_i characterize degree of burial of an atom i , which can be used to find an approximate value of SASA of an atom. In this so-called ACE-approximation of SASA is given by [30]:

$$\Delta G_{solv}^{SA} = 4\pi \sum_{i=1}^N \sigma_i (R_i + R_{probe})^2 \frac{R_i}{\alpha_i}, \quad (1.63)$$

where σ_i is the solvation parameter, R_{probe} is the radius of the solvation probe (radius of the molecule of the solvent), R_i is the radius of the i -th atom (values of these radii are equal to ρ_{0i} for HCT and OBC models). The idea of this relation follows. The Born solvation radius α_i depends on the geometry of the solute and describes the degree of burial of the atom inside the molecule. Isolated atoms have α 's relatively close to their Van der Waals radii, whereas α_i of an atom in the solute interior is on the of the molecular radius. Given that the factor $4\pi (R_i + R_{probe})^2$ represents the surface area of an isolated atom, which will be the maximum of SASA in case $\alpha_i = R_i$. When α is large and the atom is situated deeply inside the molecular interior, the value of SASA will be close to zero. Although, using these had-waiving arguments, the Eq. 1.63 makes sense in asymptotic cases, using the ratio R_i/α_i as a "scaling factor" is arguable. For instance, in Gromacs source code, the following approximation is used:

$$\Delta G_{solv}^{SA} = 4\pi \sum_{i=1}^N \sigma_i (R_i + R_{probe})^2 \left(\frac{R_i}{\alpha_i} \right)^6. \quad (1.64)$$

Values, obtained using Both Eq. 1.63 and Eq. 1.64 do not completely agree with the exact SASA described by 1.26 and 1.27. We will use the Gromacs approximation to evaluate the atomic forces because it produces results that are closer to the exact SASA values.

1.8.5 Parameters

Still model

Still's approach has a set of adjustable parameters - λ , P_1 , P_2 , P_3 , P_4 , and P_5 . Values for these are listed in [15, 21], here we summarized them in Table 1.2. The atom radii R_i that are used to estimate

Table 1.2: Still model adjustable parameters values for CHARMM19 and CHARMM22 force-fields (data are taken from [15]).

Force-field	λ	P_1	P_2	P_3	P_4	P_5
CHARMM19	0.759	0.415	0.239	1.756	10.51	1.1
CHARMM22	0.793	0.448	0.173	0.013	9.015	0.9

Born radii (Eq. 1.54) are taken from the non-bonded parameters in the force-field with one exclusion: van der Waals radius of hydrogen atoms is so small, using its force-field value in the Born radii formula can lead to numerical instabilities [14, 15, 21]. The solution for this problem is to set value of the van der Waals radius of hydrogen to 1.15\AA [14, 21] or 0.8\AA [15]. We will use the later since it is the value used when empirical parameters λ , and P_1 - P_5 were found [15].

Another important point is that the Still formalism poorly describes the Born radii when $G_{pol,i}$ is close to zero. In this case α_i becomes highly sensitive to the changes in local environment and may become negative (or can even be infinite if $G_{pol,i}$ to become zero). Although this is not described in papers, according to the CHARMM source code the large cut-off value (10^5 in CHARMM source code) for the born radii is introduced. If α is negative or larger then this value, it is set to 10^5 .

HCT/OBC models

The set of parameters for these models includes the solvation radii ρ_i , scaling factors S_i , dielectric displacement ρ_{dd} and fitting parameters α , β and γ for OBC model. As it was noted, the scaling factors were initially introduced as an atom pair based parameters, i.e. in case when ad atom i is screened by an atom j , the radius scaling multiplier S_{ij} was dependent on both atom types of atoms i and j . Later this was changed for the sake of a single atom type dependence, i.e. the dependence only on the atom type of the screening atom j . The solvation radii and screening factors are listed in, for instance, in Gromacs topologies in separate file, called “gbsa.itp” or, for some force-fields “gb.itp”. This file contains some other parameters for the non-polar part of the solvation potential that are not currently in use in Gromacs. The parameters given in these files are described in Table 1.3.

Table 1.3: Description of HCT/OBC parameters listed in Gromacs topology file).

#	Column name	Units	Description
1	Atom type	–	Force-field atom type
2	sar	nm	Radii to use for SASA evaluation
3	st	?	?
4	pi	–	SASA parameter p_i (see Eq. 1.27)
5	gbr	nm	Atom radii (ρ_{0i})
6	hct	–	Scaling factor (S_i)

As it follows from Gromacs manual [8], values in columns 2-4 are not used in Gromacs. They pertain to more elaborate surface area algorithms, such as the one from Qiu *et al.* [21] or Eq. 1.27 in particular. Column 5 contains the atomic van der Waals radii, which are used in computing the Born radii. The dielectric offset (ρ_{dd}) gets subtracted from the input van der Waals radii for the different Born radii methods, as described by Onufriev *et al.* [28]. The default value of the dielectric offset (or dielectric displacement) is $\rho_{dd} = 0.09\text{\AA}$. Column 6 is the scale factor (S_i) for the HCT and OBC models. The values are taken from the Tinker implementation of the HCT pairwise scaling method [18]. This method has been modified such that the scaling factors have been adjusted to minimize differences

between analytical surface areas and GB using the HCT algorithm. The scaling is further modified in that it is not applied pairwise as proposed by Hawkins *et al.* [18], but on a per-atom (rather than a per-pair) basis.

There are parameters α , β , γ in OBC model, which are treated as adjustable dimensionless parameters to be optimized [28]. Two sets of values for these parameters were found in original paper: $\alpha = 0.8$, $\beta = 0$, $\gamma = 2.91$ and $\alpha = 1.0$, $\beta = 0.8$, $\gamma = 4.85$.

Non-polar potential (SASA)

As it follows from Eq. 1.64, once Born radii are computed, the parameters needed for successful SASA evaluation using ACE approximation are σ_i , R_i and R_{probe} . $R_i = \rho_{0i}$ are taken to be the same radii used in GB part of the potential (column 5 in Table 1.3), R_{probe} is the solvation probe radii ($R_{probe} = 1.4\text{\AA}$ for water). Although there are some inconsistency with the reported values of the energetic factors σ_i they usually are taken identical for all atoms in GB/SA approach. In Gromacs, default values are $0.0049\text{kcal/mol}\text{\AA}^2$ for Still model and $0.0054\text{kcal/mol}\text{\AA}^2$ for the HCT/OBC approaches. Values of $0.01\text{kcal/mol}\text{\AA}^2$ [21], $0.072\text{kcal/mol}\text{\AA}^2$ [14], $0.003\text{kcal/mol}\text{\AA}^2$ and $0.003\text{kcal/mol}\text{\AA}^2$ [27] are also mentioned in literature. In rare cases dependence of the solvation energetic factors σ_i on atom types is kept [26].

1.8.6 Derivation of atomic forces

First, let us introduce denotation for the constant part of the GB potential:

$$k'_c = -\frac{k_c}{2} \left(\frac{1}{\varepsilon_{in}} - \frac{1}{\varepsilon_{out}} \right) \quad (1.65)$$

Then the solvation energy due to the Generalize Born potential becomes:

$$\frac{1}{k'_c} \Delta G_{solv}^{GB} = \sum_{i=1}^N \sum_{j=1}^N G_{ij}, \quad (1.66)$$

where

$$G_{ij} = \frac{q_i q_j}{\sqrt{r_{ij}^2 + \alpha_i \alpha_j \exp \left[-\frac{r_{ij}^2}{\gamma \alpha_i \alpha_j} \right]}} = G_{ij}(r_{ij}, \alpha_i, \alpha_j). \quad (1.67)$$

The gradient of the potential is then:

$$\frac{1}{k'_c} \nabla_l (\Delta G_{solv}^{GB}) = \sum_{i=1}^N \sum_{j=1}^N \sum_{m=1}^N \sum_{n=1}^N \frac{\partial G_{ij}}{\partial r_{mn}} \nabla_l r_{mn} + \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^N \frac{\partial G_{ij}}{\partial \alpha_k} \nabla_l \alpha_k = S_1 + S_2. \quad (1.68)$$

Given, that:

$$\nabla_l r_{mn} = \frac{\vec{r}_{mn}}{r_{mn}} (\delta_{lm} - \delta_{ln}), \quad (1.69)$$

and

$$\frac{\partial G_{ij}}{\partial r_{mn}} = \frac{\partial G_{ij}}{\partial r_{mn}} (\delta_{mi} \cdot \delta_{nj}), \quad (1.70)$$

the first term in Eq. 1.68 becomes:

$$\begin{aligned}
S_1 &= \sum_{i=1}^N \sum_{j=1}^N \sum_{m=1}^N \sum_{n=1}^N \frac{\partial G_{ij}}{\partial r_{mn}} \nabla_l r_{mn} = \sum_{i=1}^N \sum_{j=1}^N \sum_{m=1}^N \sum_{n=1}^N \frac{\partial G_{ij}}{\partial r_{mn}} \delta_{mi} \delta_{ni} \nabla_l r_{mn} = \sum_{i=1}^N \sum_{j=1}^N \frac{\partial G_{ij}}{\partial r_{ij}} \nabla_l r_{ij} = \\
&= \sum_{i=1}^N \sum_{j=1}^N \frac{\partial G_{ij}}{\partial r_{ij}} \frac{\vec{r}_{ij}}{r_{ij}} (\delta_{li} - \delta_{lj}) = \sum_{i=1}^N \sum_{j=1}^N \frac{\partial G_{ij}}{\partial r_{ij}} \frac{\vec{r}_{ij}}{r_{ij}} \delta_{li} - \sum_{i=1}^N \sum_{j=1}^N \frac{\partial G_{ij}}{\partial r_{ij}} \frac{\vec{r}_{ij}}{r_{ij}} \delta_{lj} = \\
&= \sum_{i=1}^N \frac{\partial G_{li}}{\partial r_{li}} \frac{\vec{r}_{li}}{r_{li}} - \sum_{i=1}^N \frac{\partial G_{il}}{\partial r_{il}} \frac{\vec{r}_{il}}{r_{il}}.
\end{aligned} \tag{1.71}$$

Since $G_{ij} = G_{ji}$, $r_{ij} = r_{ji}$ and $\vec{r}_{ij} = -\vec{r}_{ij}$:

$$S_1 = \sum_{i=1}^N \frac{\partial G_{il}}{\partial r_{il}} \frac{-\vec{r}_{il}}{r_{il}} - \sum_{i=1}^N \frac{\partial G_{il}}{\partial r_{il}} \frac{\vec{r}_{il}}{r_{il}} = -2 \sum_{i=1}^N \frac{\partial G_{il}}{\partial r_{il}} \frac{\vec{r}_{il}}{r_{il}}. \tag{1.72}$$

The partial derivative in Eq. 1.72 can be computed as:

$$\begin{aligned}
\frac{\partial G_{il}}{\partial r_{il}} &= q_i q_l \left(-\frac{1}{2} \right) \frac{1}{\left(r_{il}^2 + \alpha_i \alpha_l \exp \left[-\frac{r_{il}^2}{\gamma \alpha_i \alpha_l} \right] \right)^{\frac{3}{2}}} (2r_{il} + \alpha_i \alpha_l \exp \left[-\frac{r_{il}^2}{\gamma \alpha_i \alpha_l} \right]) \left(-\frac{2r_{il}}{\gamma \alpha_i \alpha_l} \right) = \\
&= q_i q_l \left(-\frac{1}{2} \right) \frac{1}{\left(r_{il}^2 + \alpha_i \alpha_l \exp \left[-\frac{r_{il}^2}{\gamma \alpha_i \alpha_l} \right] \right)^{\frac{3}{2}}} 2r_{il} \left(1 - \frac{1}{\gamma} \exp \left[-\frac{r_{il}^2}{\gamma \alpha_i \alpha_l} \right] \right)
\end{aligned} \tag{1.73}$$

Combining Eqs. 1.72 and 1.73, we get:

$$\begin{aligned}
S_1 &= -2 \sum_{i=1}^N \frac{\partial G_{il}}{\partial r_{il}} \frac{\vec{r}_{il}}{r_{il}} = -2 \sum_{i=1}^N q_i q_l \left(-\frac{1}{2} \right) \frac{1}{\left(r_{il}^2 + \alpha_i \alpha_l \exp \left[-\frac{r_{il}^2}{\gamma \alpha_i \alpha_l} \right] \right)^{\frac{3}{2}}} 2r_{il} \left(1 - \frac{1}{\gamma} \exp \left[-\frac{r_{il}^2}{\gamma \alpha_i \alpha_l} \right] \right) \frac{\vec{r}_{il}}{r_{il}} = \\
&= 2q_l \sum_{i=1}^N \frac{q_i \left(1 - \frac{1}{\gamma} \exp \left[-\frac{r_{il}^2}{\gamma \alpha_i \alpha_l} \right] \right)}{\left(r_{il}^2 + \alpha_i \alpha_l \exp \left[-\frac{r_{il}^2}{\gamma \alpha_i \alpha_l} \right] \right)^{\frac{3}{2}}} \vec{r}_{il}
\end{aligned} \tag{1.74}$$

The second term in Eq. 1.68 contains partial derivative of the Generalized Born potential over Born radii of an atom α . Since $G_{ij} = G_{ij}(r_{ij}, \alpha_i, \alpha_j)$, it can be written as:

$$\frac{\partial G_{ij}}{\partial \alpha_k} = \frac{\partial G_{ij}}{\partial \alpha_k} (\delta_{ik} + \delta_{jk}). \tag{1.75}$$

Using this relation, S_2 becomes:

$$\begin{aligned}
S_2 &= \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^N \frac{\partial G_{ij}}{\partial \alpha_k} \nabla_l \alpha_k = \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^N \frac{\partial G_{ij}}{\partial \alpha_k} (\delta_{ik} + \delta_{jk}) \nabla_l \alpha_k = \\
&= \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^N \frac{\partial G_{ij}}{\partial \alpha_k} \delta_{ik} \nabla_l \alpha_k + \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^N \frac{\partial G_{ij}}{\partial \alpha_k} \delta_{jk} \nabla_l \alpha_k = \\
&= \sum_{i=1}^N \sum_{j=1}^N \frac{\partial G_{ij}}{\partial \alpha_i} \nabla_l \alpha_i + \sum_{i=1}^N \sum_{j=1}^N \frac{\partial G_{ij}}{\partial \alpha_j} \nabla_l \alpha_j = 2 \sum_{i=1}^N \sum_{j=1}^N \frac{\partial G_{ij}}{\partial \alpha_i} \nabla_l \alpha_i
\end{aligned} \tag{1.76}$$

In latter, we used the fact that $G_{ij} = G_{ji}$. The partial derivative in Eq. 1.76 is:

$$\begin{aligned}
\frac{\partial G_{ij}}{\partial \alpha_i} &= \left(-\frac{1}{2}\right) \frac{q_i q_j}{\left(r_{ij}^2 + \alpha_i \alpha_j \exp\left[-\frac{r_{ij}^2}{\gamma \alpha_i \alpha_j}\right]\right)^{\frac{3}{2}}} \left(\alpha_j \exp\left[-\frac{r_{ij}^2}{\gamma \alpha_i \alpha_j}\right] + \alpha_i \alpha_j \exp\left[-\frac{r_{ij}^2}{\gamma \alpha_i \alpha_j}\right] \left(-\frac{r_{ij}^2}{\gamma \alpha_j}\right) \left(-\frac{1}{\alpha_i^2}\right)\right) = \\
&= \left(-\frac{1}{2}\right) \frac{q_i q_j}{\left(r_{ij}^2 + \alpha_i \alpha_j \exp\left[-\frac{r_{ij}^2}{\gamma \alpha_i \alpha_j}\right]\right)^{\frac{3}{2}}} \alpha_j \exp\left[-\frac{r_{ij}^2}{\gamma \alpha_i \alpha_j}\right] \left(1 + \alpha_i \frac{r_{ij}^2}{\gamma \alpha_i^2 \alpha_j}\right) = \\
&= \left(-\frac{1}{2}\right) \frac{q_i q_j \alpha_j \exp\left[-\frac{r_{ij}^2}{\gamma \alpha_i \alpha_j}\right]}{\left(r_{ij}^2 + \alpha_i \alpha_j \exp\left[-\frac{r_{ij}^2}{\gamma \alpha_i \alpha_j}\right]\right)^{\frac{3}{2}}} \left(1 + \frac{r_{ij}^2}{\gamma \alpha_i \alpha_j}\right)
\end{aligned} \tag{1.77}$$

The relation used to compute α_i is model-dependent. However, there are certain similarities in approaches that use pair-based form of potential (e.g. in Still, HCT and OBC models). The general form of the dependance of the Born radii α_i on the distances between atoms r_{ij} for these models can be written as:

$$\alpha_i = H(\Sigma_i) = H\left(\sum_{\substack{k=1 \\ k \neq i}}^N \alpha_{ik}\right), \tag{1.78}$$

where addition to the Born radii due to the overlap of i -th atom by j -th atom α_{ik} and function $H(\Sigma_i)$ are model-dependent. For Still's model, these are given by Eqs. 1.53 and 1.54, for HCT model – by Eqs. 1.56 and 1.58, for OBC model – by Eqs. 1.61 and 1.62. We will discuss the actual form of the α_{ik} and $H(\Sigma_i)$ later on, when the particular computational procedures for these three models will be discussed. At this point, let us derive the general part of the atomic forces, common for all three models. According to 1.78, the gradient of α_i is:

$$\nabla_l \alpha_i = \nabla_l H(\Sigma_i) = \frac{\partial H(\Sigma_i)}{\partial \Sigma_i} \left(\sum_{\substack{k=1 \\ k \neq i}}^N \frac{\partial \alpha_{ik}}{\partial r_{ik}} \delta_{il} \frac{\vec{r}_{ik}}{r_{ik}} - \frac{\partial \alpha_{il}}{\partial r_{il}} \Big|_{i \neq l} \frac{\vec{r}_{il}}{r_{il}} \right) \tag{1.79}$$

Using this equation, we can rewrite the relation for S_2 (Eq. 1.76) as follows:

$$\begin{aligned}
S_2 &= 2 \sum_{i=1}^N \sum_{j=1}^N \frac{\partial G_{ij}}{\partial \alpha_i} \nabla_l \alpha_i = 2 \sum_{i=1}^N \sum_{j=1}^N \frac{\partial G_{ij}}{\partial \alpha_i} \frac{\partial H(\Sigma_i)}{\partial \Sigma_i} \left(\sum_{\substack{k=1 \\ k \neq i}}^N \frac{\partial \alpha_{ik}}{\partial r_{ik}} \delta_{il} \frac{\vec{r}_{ik}}{r_{ik}} - \frac{\partial \alpha_{il}}{\partial r_{il}} \Big|_{i \neq l} \frac{\vec{r}_{il}}{r_{il}} \right) = \\
&= 2 \sum_{i=1}^N \sum_{j=1}^N \sum_{\substack{k=1 \\ k \neq i}}^N \frac{\partial G_{ij}}{\partial \alpha_i} \frac{\partial H(\Sigma_i)}{\partial \Sigma_i} \frac{\partial \alpha_{ik}}{\partial r_{ik}} \delta_{il} \frac{\vec{r}_{ik}}{r_{ik}} - 2 \sum_{i=1}^N \sum_{j=1}^N \frac{\partial G_{ij}}{\partial \alpha_i} \frac{\partial H(\Sigma_i)}{\partial \Sigma_i} \frac{\partial \alpha_{il}}{\partial r_{il}} \Big|_{i \neq l} \frac{\vec{r}_{il}}{r_{il}} = \\
&= 2 \sum_{j=1}^N \sum_{\substack{k=1 \\ k \neq l}}^N \frac{\partial G_{lj}}{\partial \alpha_l} \frac{\partial H(\Sigma_l)}{\partial \Sigma_l} \frac{\partial \alpha_{lk}}{\partial r_{lk}} \frac{\vec{r}_{lk}}{r_{lk}} - 2 \sum_{\substack{i=1 \\ i \neq l}}^N \sum_{j=1}^N \frac{\partial G_{ij}}{\partial \alpha_i} \frac{\partial H(\Sigma_i)}{\partial \Sigma_i} \frac{\partial \alpha_{il}}{\partial r_{il}} \frac{\vec{r}_{il}}{r_{il}} = \\
&= 2 \sum_{j=1}^N \sum_{\substack{i=1 \\ i \neq l}}^N \frac{\partial G_{lj}}{\partial \alpha_l} \frac{\partial H(\Sigma_l)}{\partial \Sigma_l} \frac{\partial \alpha_{li}}{\partial r_{li}} \frac{\vec{r}_{li}}{r_{li}} - 2 \sum_{\substack{i=1 \\ i \neq l}}^N \sum_{j=1}^N \frac{\partial G_{ij}}{\partial \alpha_i} \frac{\partial H(\Sigma_i)}{\partial \Sigma_i} \frac{\partial \alpha_{il}}{\partial r_{il}} \frac{\vec{r}_{il}}{r_{il}} = \\
&= -2 \sum_{\substack{i=1 \\ i \neq l}}^N \sum_{j=1}^N \frac{\partial G_{lj}}{\partial \alpha_l} \frac{\partial H(\Sigma_l)}{\partial \Sigma_l} \frac{\partial \alpha_{li}}{\partial r_{li}} \frac{\vec{r}_{il}}{r_{il}} - 2 \sum_{\substack{i=1 \\ i \neq l}}^N \sum_{j=1}^N \frac{\partial G_{ij}}{\partial \alpha_i} \frac{\partial H(\Sigma_i)}{\partial \Sigma_i} \frac{\partial \alpha_{il}}{\partial r_{il}} \frac{\vec{r}_{il}}{r_{il}} = \\
&= -2 \sum_{\substack{i=1 \\ i \neq l}}^N \left(\frac{\partial H(\Sigma_l)}{\partial \Sigma_l} \frac{\partial \alpha_{li}}{\partial r_{li}} \frac{\vec{r}_{il}}{r_{il}} \sum_{j=1}^N \frac{\partial G_{lj}}{\partial \alpha_l} + \frac{\partial H(\Sigma_i)}{\partial \Sigma_i} \frac{\partial \alpha_{il}}{\partial r_{il}} \frac{\vec{r}_{il}}{r_{il}} \sum_{j=1}^N \frac{\partial G_{ij}}{\partial \alpha_i} \right) = \\
&= -2 \sum_{\substack{i=1 \\ i \neq l}}^N \left(\frac{\partial H(\Sigma_l)}{\partial \Sigma_l} \frac{\partial \alpha_{li}}{\partial r_{li}} \sum_{j=1}^N \frac{\partial G_{lj}}{\partial \alpha_l} + \frac{\partial H(\Sigma_i)}{\partial \Sigma_i} \frac{\partial \alpha_{il}}{\partial r_{il}} \sum_{j=1}^N \frac{\partial G_{ij}}{\partial \alpha_i} \right) \frac{\vec{r}_{il}}{r_{il}}
\end{aligned} \tag{1.80}$$

Combining Eqs. 1.77 and 1.80, we get:

$$\begin{aligned}
S_2 &= -2 \sum_{\substack{i=1 \\ i \neq l}}^N \left(\frac{\partial H(\Sigma_l)}{\partial \Sigma_l} \frac{\partial \alpha_{li}}{\partial r_{li}} \sum_{j=1}^N \left(-\frac{1}{2} \right) \frac{q_l q_j \alpha_j \exp \left[-\frac{r_{lj}^2}{\gamma \alpha_l \alpha_j} \right]}{\left(r_{lj}^2 + \alpha_l \alpha_j \exp \left[-\frac{r_{lj}^2}{\gamma \alpha_l \alpha_j} \right] \right)^{\frac{3}{2}}} \left(1 + \frac{r_{lj}^2}{\gamma \alpha_l \alpha_j} \right) + \right. \\
&\quad \left. + \frac{\partial H(\Sigma_i)}{\partial \Sigma_i} \frac{\partial \alpha_{il}}{\partial r_{il}} \sum_{j=1}^N \left(-\frac{1}{2} \right) \frac{q_i q_j \alpha_j \exp \left[-\frac{r_{ij}^2}{\gamma \alpha_i \alpha_j} \right]}{\left(r_{ij}^2 + \alpha_i \alpha_j \exp \left[-\frac{r_{ij}^2}{\gamma \alpha_i \alpha_j} \right] \right)^{\frac{3}{2}}} \left(1 + \frac{r_{ij}^2}{\gamma \alpha_i \alpha_j} \right) \right) \frac{\vec{r}_{il}}{r_{il}}
\end{aligned} \tag{1.81}$$

$$\begin{aligned}
S_2 = & \sum_{\substack{i=1 \\ i \neq l}}^N \left(\frac{\partial H(\Sigma_l)}{\partial \Sigma_l} \frac{\partial \alpha_{li}}{\partial r_{li}} \left\{ q_l \sum_{j=1}^N \frac{q_j \alpha_j \exp \left[-\frac{r_{lj}^2}{\gamma \alpha_l \alpha_j} \right]}{\left(r_{lj}^2 + \alpha_l \alpha_j \exp \left[-\frac{r_{lj}^2}{\gamma \alpha_l \alpha_j} \right] \right)^{\frac{3}{2}}} \left(1 + \frac{r_{lj}^2}{\gamma \alpha_l \alpha_j} \right) \right\} + \right. \\
& \left. + \frac{\partial H(\Sigma_i)}{\partial \Sigma_i} \frac{\partial \alpha_{il}}{\partial r_{il}} \left\{ q_i \sum_{j=1}^N \frac{q_j \alpha_j \exp \left[-\frac{r_{ij}^2}{\gamma \alpha_i \alpha_j} \right]}{\left(r_{ij}^2 + \alpha_i \alpha_j \exp \left[-\frac{r_{ij}^2}{\gamma \alpha_i \alpha_j} \right] \right)^{\frac{3}{2}}} \left(1 + \frac{r_{ij}^2}{\gamma \alpha_i \alpha_j} \right) \right\} \right) \frac{\vec{r}_{il}}{r_{il}}
\end{aligned} \tag{1.82}$$

Finally, the general formula to compute the force acting on the atom due to pair-base Generalized Born potential will be:

$$\begin{aligned}
\frac{1}{k'_c} \vec{f}_l^{GB} = & -\frac{1}{k'_c} \nabla_l \Delta G = -2q_l \sum_{i=1}^N \frac{q_i \left(1 - \frac{1}{\gamma} \exp \left[-\frac{r_{il}^2}{\gamma \alpha_i \alpha_l} \right] \right)}{\left(r_{il}^2 + \alpha_i \alpha_l \exp \left[-\frac{r_{il}^2}{\gamma \alpha_i \alpha_l} \right] \right)^{\frac{3}{2}}} \vec{r}_{il} - \\
& - \sum_{\substack{i=1 \\ i \neq l}}^N \left(\frac{\partial H(\Sigma_l)}{\partial \Sigma_l} \frac{\partial \alpha_{li}}{\partial r_{li}} \left\{ q_l \sum_{j=1}^N \frac{q_j \alpha_j \exp \left[-\frac{r_{lj}^2}{\gamma \alpha_l \alpha_j} \right]}{\left(r_{lj}^2 + \alpha_l \alpha_j \exp \left[-\frac{r_{lj}^2}{\gamma \alpha_l \alpha_j} \right] \right)^{\frac{3}{2}}} \left(1 + \frac{r_{lj}^2}{\gamma \alpha_l \alpha_j} \right) \right\} + \right. \\
& \left. + \frac{\partial H(\Sigma_i)}{\partial \Sigma_i} \frac{\partial \alpha_{il}}{\partial r_{il}} \left\{ q_i \sum_{j=1}^N \frac{q_j \alpha_j \exp \left[-\frac{r_{ij}^2}{\gamma \alpha_i \alpha_j} \right]}{\left(r_{ij}^2 + \alpha_i \alpha_j \exp \left[-\frac{r_{ij}^2}{\gamma \alpha_i \alpha_j} \right] \right)^{\frac{3}{2}}} \left(1 + \frac{r_{ij}^2}{\gamma \alpha_i \alpha_j} \right) \right\} \right) \frac{\vec{r}_{il}}{r_{il}}
\end{aligned} \tag{1.83}$$

Note, that this equation has two summations folded in two terms and that the internal summations (in curly brackets) are identical in both cases. Let us introduce the following denotations:

$$S_i^\alpha = q_i \sum_{j=1}^N \frac{q_j \alpha_j \exp \left[-\frac{r_{ij}^2}{\gamma \alpha_i \alpha_j} \right]}{\left(r_{ij}^2 + \alpha_i \alpha_j \exp \left[-\frac{r_{ij}^2}{\gamma \alpha_i \alpha_j} \right] \right)^{\frac{3}{2}}} \left(1 + \frac{r_{ij}^2}{\gamma \alpha_i \alpha_j} \right), \tag{1.84}$$

and

$$\vec{S}_i^r = -2q_l \sum_{i=1}^N \frac{q_i \left(1 - \frac{1}{\gamma} \exp \left[-\frac{r_{il}^2}{\gamma \alpha_i \alpha_l} \right] \right)}{\left(r_{il}^2 + \alpha_i \alpha_l \exp \left[-\frac{r_{il}^2}{\gamma \alpha_i \alpha_l} \right] \right)^{\frac{3}{2}}} \vec{r}_{il}. \tag{1.85}$$

Then, the atomic force can be computed as:

$$\vec{f}_l^{GB} = k'_c \vec{S}_i^r - k'_c \sum_{\substack{i=1 \\ i \neq l}}^N \left(\frac{\partial H(\Sigma_l)}{\partial \Sigma_l} \frac{\partial \alpha_{li}}{\partial r_{li}} S_l^\alpha + \frac{\partial H(\Sigma_i)}{\partial \Sigma_i} \frac{\partial \alpha_{il}}{\partial r_{il}} S_i^\alpha \right) \frac{\vec{r}_{il}}{r_{il}} \tag{1.86}$$

Still model

Denote

$$C_i^{\lambda, P_1} = \frac{1}{\lambda} \left(-\frac{k_c}{2R_i} \right) + P_1 \left(\frac{k_c}{2R_i^2} \right) \quad (1.87)$$

and

$$G_{ij} = P \frac{V_j}{r_{ij}^4} [CCF]_{ij} \quad (1.88)$$

where $P = P_2, P_3$ or P_4 and $[CCF]_{ij} = 1, 1$ or CCF_{ij} for bonded atoms, atoms connected through two covalent bonds (angle) or non-bonded atoms respectively. Then:

$$G_{pol,i} = C_i^{\lambda, P_1} + \sum_{j \neq i}^N G_{ij} \quad (1.89)$$

If we denote $g_{ij} = \frac{G_{ij}}{V_j} = \frac{G_{ji}}{V_i} = g_{ji}$, then:

$$\nabla_l G_{pol,i} = \nabla_l \sum_{\substack{j=1 \\ j \neq i}}^N G_{ij} = \nabla_l \sum_{\substack{j=1 \\ j \neq i}}^N g_{ij} V_j = \sum_{\substack{j=1 \\ j \neq i}}^N V_j \nabla_l g_{ij} = \begin{cases} \sum_{\substack{j=1 \\ j \neq l}}^N V_j \nabla_l g_{lj} & \text{when } i = l \\ V_l \nabla_l g_{il} & \text{when } i \neq l \end{cases} \quad (1.90)$$

where, using Eq. 1.88, we find:

$$\vec{\sigma}_{lj} \equiv \nabla_l g_{lj} = P \nabla_l \left(\frac{[CCF]_{lj}}{r_{lj}^4} \right) = P \left(-\frac{4}{r_{lj}^6} \vec{r}_{lj} [CCF]_{lj} + \frac{1}{r_{lj}^4} \nabla_l ([CCF]_{lj}) \right) \quad (1.91)$$

From Eq. 1.55 it follows that:

$$\nabla_l [CCF]_{lj} = \begin{cases} 0, & \text{when } \left(\frac{r_{lj}}{R_l + R_j} \right)^2 \text{ or atoms } l \text{ and } j \text{ are 1-2 or 1-3 bonded} \\ \nabla_l \left(\frac{1}{4} \left(1 - \cos \left[\pi P_5 \left(\frac{r_{lj}}{R_l + R_j} \right)^2 \right] \right)^2 \right) & \text{otherwise} \end{cases} \quad (1.92)$$

Expanding second case, we get:

$$\begin{aligned} \nabla_l \left(\frac{1}{4} \left(1 - \cos \left[\pi P_5 \left(\frac{r_{lj}}{R_l + R_j} \right)^2 \right] \right)^2 \right) &= \\ = \frac{1}{4} 2 \left(1 - \cos \left[\pi P_5 \left(\frac{r_{lj}}{R_l + R_j} \right)^2 \right] \right) \sin \left[\pi P_5 \left(\frac{r_{lj}}{R_l + R_j} \right)^2 \right] \pi P_5 \left(\frac{r_{lj}}{R_l + R_j} \right)^2 \frac{\vec{r}_{lj}}{r_{lj}} \end{aligned} \quad (1.93)$$

Using denotation:

$$\theta_{lj} = \pi P_5 \left(\frac{r_{lj}}{R_l + R_j} \right)^2 \quad (1.94)$$

Eq. 1.92 becomes:

$$\nabla_l [CCF]_{lj} = \begin{cases} 0, & \text{when } \theta_{lj} > \pi \text{ or atoms } l \text{ and } j \text{ are 1-2 or 1-3 bonded} \\ (1 - \cos \theta_{lj}) \sin \theta_{lj} \theta_{lj} \frac{\vec{r}_{lj}}{r_{lj}}, & \text{otherwise} \end{cases} \quad (1.95)$$

Combining Eqs. 1.88, 1.91 and 1.95, we get:

$$\vec{\sigma}_{lj} = \begin{cases} -\frac{4P_2}{r_{lj}^6} \vec{r}_{lj}, & \text{when atoms } l \text{ and } j \text{ are bonded} \\ -\frac{4P_3}{r_{lj}^6} \vec{r}_{lj}, & \text{when atoms } l \text{ and } j \text{ are 1-3 bonded} \\ -\frac{4P_4}{r_{lj}^6} \vec{r}_{lj}, & \text{when atoms } l \text{ and } j \text{ are not bonded} \\ & \text{and } \theta_{lj} > \pi \\ -\frac{4P_4}{r_{lj}^6} \left[(1 - \cos \theta_{lj})^2 - (1 - \cos \theta_{lj}) \sin \theta_{lj} \theta_{lj} \right] \vec{r}_{lj}, & \text{when atoms } l \text{ and } j \text{ are not bonded} \\ & \text{and } \theta_{lj} \leq \pi \end{cases} \quad (1.96)$$

Note, that:

$$\nabla_l \alpha_i = \begin{cases} \frac{2\alpha_i^2}{k_c} V_l \sigma_{li} \vec{r}_{li}, & \text{when } i \neq l \\ \frac{2\alpha_l^2}{k_c} \sum_{j=1, j \neq l}^N V_j \sigma_{lj} \vec{r}_{lj}, & \text{when } i = l \end{cases} \quad (1.97)$$

HCT model

Using Eq. 1.58, the equations for the function $H(\Sigma_i)$ and terms α_{ij} (see Eq. 1.78) can be written as:

$$\alpha_i = H(\Sigma_i) = \frac{1}{\frac{1}{\rho_i} - \frac{1}{2}\Sigma_i}, \quad (1.98)$$

and:

$$\alpha_{ij} = \frac{1}{L_{ij}} - \frac{1}{U_{ij}} + \frac{r_{ij}}{4} \left(\frac{1}{U_{ij}^2} - \frac{1}{L_{ij}^2} \right) + \frac{1}{2r_{ij}} \ln \frac{L_{ij}}{U_{ij}} + \frac{(S_j \rho_j)^2}{4r_{ij}} \left(\frac{1}{L_{ij}^2} - \frac{1}{U_{ij}^2} \right) \quad (1.99)$$

Then, for HCT model:

$$\frac{\partial H(\Sigma_i)}{\partial \Sigma_i} = \frac{\partial}{\partial \Sigma_i} \left(\frac{1}{\frac{1}{\rho_i} - \frac{1}{2}\Sigma_i} \right) = -\frac{1}{\left(\frac{1}{\rho_i} - \frac{1}{2}\Sigma_i \right)^2} \left(-\frac{1}{2} \right) = \frac{1}{2} \alpha_i^2, \quad (1.100)$$

and:

$$\begin{aligned} \frac{\partial \alpha_{ij}}{\partial r_{ij}} &= -\frac{1}{L_{ij}^2} \frac{\partial L_{ij}}{\partial r_{ij}} + \frac{1}{U_{ij}^2} \frac{\partial U_{ij}}{\partial r_{ij}} + \frac{1}{4} \left[\left(-\frac{2}{U_{ij}^3} \frac{\partial U_{ij}}{\partial r_{ij}} + \frac{2}{L_{ij}^3} \frac{\partial L_{ij}}{\partial r_{ij}} \right) \left(r_{ij} - \frac{S_j^2 \rho_j^2}{r_{ij}} \right) + \right. \\ &\quad \left. + \left(\frac{1}{U_{ij}^2} - \frac{1}{L_{ij}^2} \right) \left(1 + \frac{S_j^2 \rho_j^2}{r_{ij}^2} \right) \right] - \frac{1}{2r_{ij}^2} \ln \frac{L_{ij}}{U_{ij}} + \frac{1}{2r_{ij}} \frac{U_{ij}}{L_{ij}} \left(\frac{1}{U_{ij}} \frac{\partial L_{ij}}{\partial r_{ij}} - \frac{L_{ij}}{U_{ij}^2} \frac{\partial U_{ij}}{\partial r_{ij}} \right) = \\ &= -\frac{1}{L_{ij}^2} \frac{\partial L_{ij}}{\partial r_{ij}} + \frac{1}{U_{ij}^2} \frac{\partial U_{ij}}{\partial r_{ij}} + \frac{1}{2} \left(-\frac{1}{U_{ij}^3} \frac{\partial U_{ij}}{\partial r_{ij}} + \frac{1}{L_{ij}^3} \frac{\partial L_{ij}}{\partial r_{ij}} \right) \left(r_{ij} - \frac{S_j^2 \rho_j^2}{r_{ij}} \right) + \\ &\quad + \frac{1}{4} \left(\frac{1}{U_{ij}^2} - \frac{1}{L_{ij}^2} \right) \left(1 + \frac{S_j^2 \rho_j^2}{r_{ij}^2} \right) - \frac{1}{2r_{ij}^2} \ln \frac{L_{ij}}{U_{ij}} + \frac{1}{2r_{ij}} \left(\frac{1}{L_{ij}} \frac{\partial L_{ij}}{\partial r_{ij}} - \frac{1}{U_{ij}} \frac{\partial U_{ij}}{\partial r_{ij}} \right), \end{aligned} \quad (1.101)$$

where, using Eqs. 1.59 and 1.60:

$$\frac{\partial L_{ij}}{\partial r_{ij}} = \begin{cases} 0 & \text{if } r_{ij} - S_j \rho_j \leq \rho_i \\ 1 & \text{if } \rho_i \leq r_{ij} - S_j \rho_j \end{cases} \quad (1.102)$$

and

$$\frac{\partial U_{ij}}{\partial r_{ij}} = \begin{cases} 0 & \text{if } r_{ij} + S_j \rho_j \leq \rho_i \\ 1 & \text{if } \rho_i \leq r_{ij} - S_j \rho_j \end{cases} \quad (1.103)$$

OBC model

As it follows from Eqs. 1.61 and 1.62:

$$\alpha_i = H(\Sigma_i) = \left(\frac{1}{\rho_i} - \frac{1}{\rho_{0i}} \tanh \left[\alpha \rho_i \frac{1}{2} \Sigma_i - \beta \left(\rho_i \frac{1}{2} \Sigma_i \right)^2 + \gamma \left(\rho_i \frac{1}{2} \Sigma_i \right)^3 \right] \right)^{-1}. \quad (1.104)$$

Then:

$$\begin{aligned} \frac{\partial H(\Sigma_i)}{\partial \Sigma_i} &= - \left(\frac{1}{\rho_i} - \frac{1}{\rho_{0i}} \tanh \left[\alpha \rho_i \frac{1}{2} \Sigma_i - \beta \left(\rho_i \frac{1}{2} \Sigma_i \right)^2 + \gamma \left(\rho_i \frac{1}{2} \Sigma_i \right)^3 \right] \right)^{-2} \cdot \left(-\frac{1}{\rho_i} \right) \cdot \\ &\cdot \left(1 - \tanh^2 \left[\alpha \rho_i \frac{1}{2} \Sigma_i - \beta \left(\rho_i \frac{1}{2} \Sigma_i \right)^2 + \gamma \left(\rho_i \frac{1}{2} \Sigma_i \right)^3 \right] \right) \left(\alpha \frac{1}{2} \rho_i - \beta \frac{1}{2} \Sigma_i \rho_i^2 + \gamma \frac{3}{8} \Sigma_i^2 \rho_i^3 \right) = \\ &= \alpha_i^2 \frac{\rho_i}{\rho_{0i}} \left(1 - \tanh^2 \left[\alpha \rho_i \frac{1}{2} \Sigma_i - \beta \left(\rho_i \frac{1}{2} \Sigma_i \right)^2 + \gamma \left(\rho_i \frac{1}{2} \Sigma_i \right)^3 \right] \right) \cdot \\ &\cdot \left(\alpha \frac{1}{2} - \beta \left(\frac{1}{2} \Sigma_i \rho_i \right) + \gamma \frac{3}{2} \left(\frac{1}{2} \Sigma_i \rho_i \right)^2 \right), \end{aligned} \quad (1.105)$$

Non-polar (SASA) addition to atomic forces

The non-polar contribution to atomic forces can be derived from Eq. 1.64 as follows:

$$\begin{aligned} \nabla_l(\Delta G_{sol}^{SA}) &= 4\pi \sum_{i=1}^N \sigma_i (R_i + R_{probe})^2 (-6) \frac{R_i^6}{\alpha_i^7} \frac{\partial H(\Sigma_i)}{\partial \Sigma_i} \left(\sum_{\substack{k=1 \\ k \neq i}}^N \frac{\partial \alpha_{ik}}{\partial r_{ik}} \delta_{il} \frac{\vec{r}_{ik}}{r_{ik}} - \frac{\partial \alpha_{il}}{\partial r_{il}} \Big|_{i \neq l} \frac{\vec{r}_{ik}}{r_{ik}} \right) = \\ &= 4\pi \sigma_l (R_l + R_{probe})^2 (-6) \frac{R_l^6}{\alpha_l^7} \frac{\partial H(\Sigma_l)}{\partial \Sigma_l} \sum_{\substack{k=1 \\ k \neq l}}^N \frac{\partial \alpha_{lk}}{\partial r_{lk}} \frac{\vec{r}_{lk}}{r_{lk}} - \\ &- 4\pi \sum_{\substack{i=1 \\ i \neq l}}^N \sigma_i (R_i + R_{probe})^2 (-6) \frac{R_i^6}{\alpha_i^7} \frac{\partial H(\Sigma_i)}{\partial \Sigma_i} \frac{\partial \alpha_{il}}{\partial r_{il}} \frac{\vec{r}_{ik}}{r_{ik}} = -\vec{f}_l^{SA}. \end{aligned} \quad (1.106)$$

Changing summation indeces from k to i in the first term, we get:

$$\vec{f}_l^{SA} = -12\pi \sum_{\substack{i=1 \\ i \neq l}}^N \left(\sigma_l (R_l + R_{probe})^2 \frac{R_l^6}{\alpha_l^7} \frac{\partial H(\Sigma_l)}{\partial \Sigma_l} \frac{\partial \alpha_{li}}{\partial r_{li}} + \sigma_i (R_i + R_{probe})^2 \frac{R_i^6}{\alpha_i^7} \frac{\partial H(\Sigma_i)}{\partial \Sigma_i} \frac{\partial \alpha_{il}}{\partial r_{il}} \right) \frac{\vec{r}_{il}}{r_{il}}. \quad (1.107)$$

During the computational process, pre-computed values for the pre-factors can be used:

$$P_i^{SA} = 12\pi\sigma_i (R_i + R_{probe})^2 R_i^6 \quad (1.108)$$

Using Eqs. 1.86, 1.107 and 1.108, we can rewrite the general formula for GB atomic forces to obtain equation for forces in GB/SA approach:

$$\begin{aligned} \vec{f}_l^{GB/SA} &= \vec{f}_l^{GB} + \vec{f}_l^{SA} = k'_c \vec{S}_i^r - \sum_{\substack{i=1 \\ i \neq l}}^N \left(k'_c \frac{\partial H(\Sigma_l)}{\partial \Sigma_l} \frac{\partial \alpha_{li}}{\partial r_{li}} S_l^\alpha + k'_c \frac{\partial H(\Sigma_i)}{\partial \Sigma_i} \frac{\partial \alpha_{il}}{\partial r_{il}} S_i^\alpha \right) \frac{\vec{r}_{il}}{r_{il}} - \\ &\quad - \sum_{\substack{i=1 \\ i \neq l}}^N \left(\frac{P_l^{SA}}{\alpha_l^7} \frac{\partial H(\Sigma_l)}{\partial \Sigma_l} \frac{\partial \alpha_{li}}{\partial r_{li}} + \frac{P_i^{SA}}{\alpha_i^7} \frac{\partial H(\Sigma_i)}{\partial \Sigma_i} \frac{\partial \alpha_{il}}{\partial r_{il}} \right) \frac{\vec{r}_{il}}{r_{il}} = \\ &= k'_c \vec{S}_i^r - \sum_{\substack{i=1 \\ i \neq l}}^N \left[\frac{\partial H(\Sigma_l)}{\partial \Sigma_l} \frac{\partial \alpha_{li}}{\partial r_{li}} \left(k'_c S_l^\alpha + \frac{P_l^{SA}}{\alpha_l^7} \right) + \left(k'_c S_i^\alpha + \frac{P_i^{SA}}{\alpha_i^7} \right) \right] \frac{\vec{r}_{il}}{r_{il}}. \end{aligned} \quad (1.109)$$

1.8.7 Computational algorithm

In general, the computational algorithm for force evaluation can be summarized in following three steps:

Step 1: At the first step of the computational algorithm Born radii for all atoms are computed. This is done using Eqs. 1.53-1.55 (Still model), Eqs. 1.58-1.60 (HCT model) or Eqs. 1.61, 1.62 and 1.59-1.60 (OBC model). This step is model dependent and includes evaluation of α_{ij} for all atom pairs i - j , finding the sum Σ_i and computing $H(\Sigma_i)$ to find the final value of α_i (see general Eq. 1.78). Due to the fact, that in OBC approach, $\partial H(\Sigma_i)/\partial \Sigma_i$ depends on value of Σ_i itself, not only on α_i , the value of $\partial H(\Sigma_i)/\partial \Sigma_i$ should be computed and saved during this step for this model as well (Eq. 1.105).

Step 2: This step contains two parts that can be done simultaneously: evaluation of summations S_i^α (Eq. 1.84) and vector-sums S_i^r (Eq. 1.85). Since these relations are model-independent, the computational procedures can be unified for all three GB approaches described here.

Step 3: On step 3, the values obtained on previous steps are combined together to produce the atomic forces according to the Eq. 1.86 (GB model) or Eq. 1.109 (GB/SA model).

1.9 Long-range electrostatics

1.9.1 Ewald summation

We start from Coulomb law:

$$E_{Col} = \frac{1}{4\pi\epsilon_0} \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N \frac{q_i q_j}{r_{ij}}. \quad (1.110)$$

Now suppose that charges are subject to the periodic boundary condition, which implies that each of the charges have an infinite number of replicas or images. Given the periodic boundary conditions, one

have to take into account an electrostatic effect of these imagery charges:

$$E_{Col} = \frac{1}{2} \frac{1}{4\pi\epsilon_0} \sum_{\vec{n} \in \mathbb{Z}^3} \sum_{i=1}^N \sum_{j=1}^{N'} \frac{q_i q_j}{|\vec{r}_{ij} + \vec{n}L|}, \quad (1.111)$$

where $\vec{n}L$ takes into account the summation over all replicas, prime ($'$) indicates that the term $i = j$ has to be excluded for the supercell (when $n = 0$). As one can see, this potential is slowly decaying with distance ($\sim 1/r$) and the sum converges very slowly. The Ewald method allows one to evaluate E_{Col} by transforming it into summation that converges rapidly [31–33].

Bibliography

- [1] A. D. MacKerell, Jr., D. Bashford, M. Bellott, R. L. Dunbrack, Jr., J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, III, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, and M. Karplus, "All-atom empirical potential for molecular modeling and dynamics studies of proteins," *J. Phys. Chem. B*, vol. 102, no. 18, pp. 3586–3616, 1998.
- [2] A. D. MacKerell, "Empirical force fields for biological macromolecules: Overview and issues," *J. Comput. Chem.*, vol. 25, no. 13, pp. 1584–1604, 2004.
- [3] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, "CHARMM: A program for macromolecular energy, minimization, and dynamics calculations," *J. Comput. Chem.*, vol. 4, no. 2, pp. 187–217, 1983.
- [4] B. R. Brooks, C. L. Brooks, III, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caffisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus, "CHARMM: The biomolecular simulation program," *J. Comput. Chem.*, vol. 30, no. 10, pp. 1545–1614, 2009.
- [5] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, and K. Schulten, "Scalable molecular dynamics with NAMD," *J. Comput. Chem.*, vol. 26, pp. 1781–1802, 2005.
- [6] B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl, "GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation," *J. Chem. Theory Comput.*, vol. 4, pp. 435–447, 2008.
- [7] D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang, and R. J. Woods, "The Amber biomolecular simulation programs," *J. Comput. Chem.*, vol. 26, no. 16, pp. 1668–1688, 2005.
- [8] D. van der Spoel, E. Lindahl, B. Hess, C. Kutzner, A. R. van Buuren, E. Apol, P. J. Meulenhoff, D. P. Tieleman, A. L. T. M. Sijbers, K. A. Feenstra, R. van Drunen, and H. J. C. Berendsen, *GROMACS User Manual*. The GROMACS development team, 4.0 ed., 2009.
- [9] A. D. MacKerell, M. Feig, and C. L. Brooks, III, "Improved treatment of the protein backbone in empirical force fields," *JACS*, vol. 126, no. 3, pp. 698–699, 2004.
- [10] A. D. MacKerell, M. Feig, and C. L. Brooks, III, "Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations," *J. Comput. Chem.*, vol. 25, no. 11, pp. 1400–1415, 2004.

- [11] J. Chen, C. L. Brooks, III, and J. Khandogin, "Recent advances in implicit solvent-based methods for biomolecular simulations," *Cur. Opin. Struct. Biol.*, vol. 18, no. 2, pp. 140–148, 2008.
- [12] P. Ferrara, J. Apostolakis, and A. Caffisch, "Evaluation of a fast implicit solvent model for molecular dynamics simulations," *Proteins*, vol. 46, no. 1, pp. 24–33, 2002.
- [13] F. Fraternali and W. F. van Gunsteren, "An efficient mean solvation force model for use in molecular dynamics simulations of proteins in aqueous solution," *J. Mol. Biol.*, vol. 256, no. 5, pp. 939–948, 1996.
- [14] W. C. Still, A. Tempczyk, R. C. Hawley, and T. Hendrickson, "Semianalytical treatment of solvation for molecular mechanics and dynamics," *JACS*, vol. 112, no. 16, pp. 6127–6129, 1990.
- [15] B. N. Dominy and C. L. Brooks, III, "Development of a Generalized Born model parametrization for proteins and nucleic acids," *J. Phys. Chem. B*, vol. 103, no. 18, pp. 3765–3773, 1999.
- [16] M. S. Lee, F. R. Salsbury, Jr., and C. L. Brooks, III, "Novel Generalized Born methods," *J. Chem. Phys.*, vol. 116, no. 24, pp. 10606–10614, 2002.
- [17] M. S. Lee, M. Feig, F. R. Salsbury, and C. L. Brooks, III, "New analytic approximation to the standard molecular volume definition and its application to generalized Born calculations," *J. Comput. Chem.*, vol. 24, no. 11, pp. 1348–1356, 2003.
- [18] G. D. Hawkins, C. J. Cramer, and D. G. Truhlar, "Parametrized models of aqueous free energies of solvation based on pairwise descreening of solute atomic charges from a dielectric medium," *J. Phys. Chem.*, vol. 100, no. 51, pp. 19824–19839, 1996.
- [19] A. Onufriev, D. Bashford, and D. A. Case, "Modification of the Generalized Born model suitable for macromolecules," *J. Phys. Chem. B*, vol. 104, no. 15, pp. 3712–3720, 2000.
- [20] T. Lazaridis and M. Karplus, "Effective energy function for proteins in solution," *Proteins*, vol. 35, no. 2, pp. 133–152, 1999.
- [21] D. Qiu, P. S. Shenkin, F. P. Hollinger, and W. C. Still, "The GB/SA continuum model for solvation. a fast analytical method for the calculation of approximate Born radii," *J. Phys. Chem. A*, vol. 101, no. 16, pp. 3005–3014, 1997.
- [22] D. Eisenberg and A. D. McLachlan, "Solvation energy in protein folding and binding," *Nature*, vol. 319, pp. 199–203, 1986.
- [23] W. Im, M. S. Lee, and C. L. Brooks, "Generalized Born model with a simple smoothing function," *J. Comput. Chem.*, vol. 24, no. 14, pp. 1691–1702, 2003.
- [24] W. Im, M. Feig, and C. L. Brooks, III, "An implicit membrane Generalized Born theory for the study of structure, stability, and interactions of membrane proteins," *Biophys. J.*, vol. 85, no. 5, pp. 2900–2918, 2003.
- [25] M. K. Gilson and B. Honig, "The inclusion of electrostatic hydration energies in molecular mechanics calculations," *J. Comput. Aid Mol. Des.*, vol. 5, no. 1, pp. 5–20, 1991.
- [26] G. D. Hawkins, C. J. Cramer, and D. G. Truhlar, "Pairwise solute descreening of solute charges from a dielectric medium," *Chem. Phys. Lett.*, vol. 246, no. 1-2, pp. 122–129, 1995.
- [27] M. Schaefer and C. Froemmel, "A precise analytical method for calculating the electrostatic energy of macromolecules in aqueous solution," *J. Mol. Biol.*, vol. 216, no. 4, pp. 1045–1066, 1990.
- [28] A. Onufriev, D. Bashford, and D. A. Case, "Exploring protein native states and large-scale conformational changes with a modified Generalized Born model," *Proteins*, vol. 55, no. 2, pp. 383–394, 2004.

- [29] W. Hasel, T. F. Hendrickson, and W. Still, "A rapid approximation to the solvent accessible surface areas of atoms," *Tetrahedron Comput. Methodol.*, vol. 1, no. 2, pp. 103–116, 1988.
- [30] M. Schaefer, C. Bartels, and M. Karplus, "Solution conformations and thermodynamics of structured peptides: molecular dynamics simulation with an implicit solvation model," *J. Mol. Biol.*, vol. 284, no. 3, pp. 835–848, 1998.
- [31] P. P. Ewald, "Die Berechnung optischer und elektrostatischer Gitterpotentiale," *Ann. Physik (Leipzig)*, vol. 369, no. 3, pp. 253–287, 1921.
- [32] H. Lee and W. Cai, "Ewald summation for Coulomb interactions in a periodic supercell," *Published online*, 2009.
- [33] C. Sagui and T. A. Darden, "Molecular dynamics simulations of biomolecules: long-range electrostatic effects," *Annu. Rev. Biophys. Biomol. Struct.*, vol. 28, pp. 155–179, 1999.