

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Energy Minimization

Budhayash Gautam

Abstract

The energetic state of a protein is one of the most important representative parameters of its stability. The energy of a protein can be defined as a function of its atomic coordinates. This energy function consists of several components: 1. Bond energy and angle energy, representative of the covalent bonds, bond angles. 2. Dihedral energy, due to the dihedral angles. 3. A van der Waals term (also called Leonard-Jones potential) to ensure that atoms do not have steric clashes. 4. Electrostatic energy accounting for the Coulomb's Law in protein structure, i.e. the long-range forces between charged and partially charged atoms. All these quantitative terms have been parameterized and are collectively referred to as the 'force-field', for e.g. CHARMM, AMBER, AMBERJOPLS and GROMOS. The goal of energy Minimization is to find a set of coordinates representing the minimum energy conformation for the given structure. Various algorithms have been formulated by varying the use of derivatives. Three common algorithms used for this optimization are steepest descent, conjugate gradient and Newton–Raphson. Although energy Minimization is a tool to achieve the nearest local minima, it is also an indispensable tool in correcting structural anomalies, viz. bad stereo-chemistry and short contacts. An efficient optimization protocol could be devised from these methods in conjunction with a larger space exploration algorithm, e.g. molecular dynamics.

Keywords: energy minimization, minimum energy conformation, force fields, global minimum energy, molecular dynamics simulations, molecular modeling

1. Introduction

Molecular modeling relies on the event of theoretical and computational methodologies, to model and study the behavior of molecules, from little chemical systems to big biological molecules and material assemblies. The applying fields of molecular modeling regard computational chemistry, drug design, computational biology and materials science. The fundamental computational technique to perform molecular modeling is simulation. Molecular simulation techniques need specific extra computational and code software system [1]. Most molecular modeling studies involve three stages. Within the initial stage a model is chosen to explain the intra- and inter-molecular interaction within the system. The two most typical models that are utilized in molecular modeling are quantum mechanics and molecular mechanics. These models enable the energy of any arrangement of the atoms and molecules within the system to be calculated and permit the modeler to work out how the energy of the system varies because the positions of the atoms and molecules change. The second stage of a molecular modeling study is that the

calculation itself, such as an energy Minimization, a molecular dynamics or Monte Carlo simulation, or a Conformations search. Finally, the calculation should be analyzed, not solely to calculate properties however additionally to see that it's been performed properly [2].

In molecular modeling we tend to be particularly curious about minimum points on the energy surface. Minimum energy arrangements of the atoms correspond to stable states of the system; any movement off from a minimum provides a configuration with a better energy. There is also a really sizable amount of minima on the energy surface. The minimum with the very lowest energy is known as the global energy minimum. To spot those geometries of the system that correspond to minimum points on the energy surface we tend to use a Minimization algorithm. The highest point on the pathway between two minima is of particular interest and is understood as the saddle point, with the arrangement of the atoms being the transition structure. Both minima and saddle points are stationary points on the energy surface, wherever the primary derivative of the energy function is zero with regard to all the coordinates [3].

A geographical analogy is useful thanks to illustrate several of the ideas as during this analogy minimum points correspond to the lowest of valleys. A minimum is also represented as being in an exceedingly 'long and slender valley' or 'a flat and featureless plain'. Saddle points correspond to mountain passes. Confer with consult with algorithms creating steps as 'uphill' and downhill'.

1.1 Energy minimization: a brief description about the problem

The Minimization problem can be formally stated as follows: given a function f which depends on one or more independent variables x_1, x_2, \dots, x_i , find the values of those variables where f has a minimum value. At a minimum point the first derivative of the function with respect to each of variables is zero and the second derivative are all positive:

$$\partial f / \partial x_i = 0; \partial^2 f / \partial x_i^2 > 0 \quad (1)$$

With respect to present discussion, the most important functions are the quantum mechanics or molecular mechanics energy with the variables x_i being the Cartesian or the internal co-ordinates of the atoms. It is a common practice to always perform Molecular mechanics Minimizations in Cartesian co-ordinates, in which the energy is a function of $3N$ variables; on the other hand, for quantum mechanics internal co-ordinates are often used. The least value of any function can be identified using standard calculus methods for analytical functions. But, due to the complexities of pattern of energy change with change in the coordinates, it is almost impossible for any molecular system. Therefore, the energy minima are often identified with the help of numerical methods. These methods gradually make changes to the coordinates to generate configurations having lower and lower energies until the minimum is reached [2].

Minimization algorithms can be classified into two categories: one in which we use derivatives of the energy with respect to the coordinates and second in which we do not use any derivative. Derivatives are extremely important because they have details about the shape of the energy surface and due to this efficiency to locate the minimum energy is increases drastically. For any problem, before choosing best algorithm (or algorithms), several points should be considered for e.g. the best Minimization algorithm should use least memory to generate the answer as quickly as possible. For different problems of molecular modeling, different Minimization

method is used because we do not have any Minimization method developed which could be applied to all. Any method which is developed for efficient performance with quantum mechanics may or may not be compatible for molecular mechanics because quantum mechanics deals with models having very less atoms as compare to molecular mechanics. Another point is that procedures like inversion of matrix in some Minimization methods works fine for small systems but problem arises when number of atoms increases to thousands. To calculate the number of derivatives of different Conformations and their energies, different level of performance is required for quantum mechanics than molecular mechanics. Molecular mechanics requires an algorithm that is having more number of steps; quantum mechanics has the opposite scenario. Therefore we have various methods in various popular software packages [4].

As, most of the Minimization algorithms can only identify the minimum energy point which is closest to the starting point, thus it can be stated that they only move downwards or more appropriately downhill on the energy surface. Suppose, this schematic energy surface is shown in **Figure 1**, having three starting points A, B and C to obtained the minima. The locations at which any hypothetical ball stops rolling on energy surface under gravitational force will have corresponding energy minima. But more important thing is to identify global energy minimum which can only be generated by using different starting points, which will be minimized later. Using this criterion, some of the Minimization methods can move uphill to find out energy minimum than the closest one. But not a single algorithm till date is reported for efficiently identification of the global minimum energy from a random starting point. To identify the number of different minimum energy Conformations the shape of energy surface is very useful. For example, population or number of minimum in a deep and narrow valley will be very less than population at broad minimum because it is having higher energy as the vibrational energy is more widely spaced in the minimum and so less accessible. Therefore, the global energy minimum may not be the most highly populated minimum. Thus, there may be the case that the 'functional' structure (e.g. the biologically active conformation of a drug molecule) may not belong to the global minimum, or to the most highly populated conformation, or even to a minimum energy structure at all [3].

Every Minimization algorithm has a set of initial coordinates as input. These coordinates can be generated from different sources. These can be generated using traditional experimental method like X-ray crystallography or NMR. Alternatively, these can be generated by employing a theoretical method like conformational search procedure. But for practical efficiency, both types of methods can be used combinatorially. For example, the behavior of a protein in water can be studied

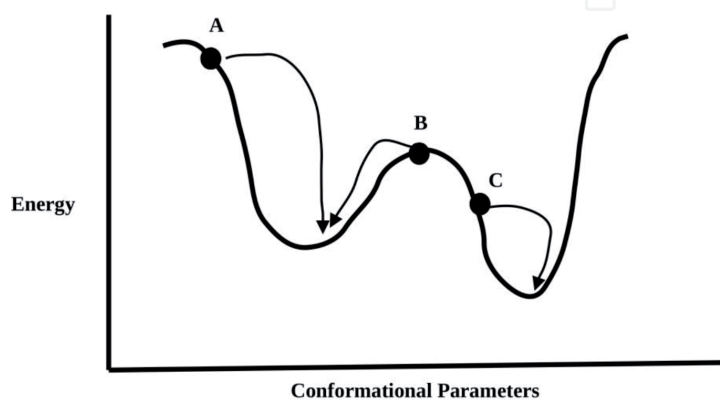


Figure 1.
A one-dimensional energy surface showing minimization methods movement downwards or downhill towards the closest energy minimum.

using its x-ray generated structure. Then place this in a solvent completely. Monte Carlo or molecular dynamics simulation can generate the atomics or Cartesian coordinates of the solvent molecules.

1.2 Derivatives

For derivatives based Minimization methods, calculation of the derivatives of the energy is performed with respect to the different variables i.e. Cartesian or internal coordinates, as the case may be. These derivatives can be generated using either analytical or numerical procedures but derivatives obtained through analytical procedure are more preferred because these can be generated more readily and these are more exact. Although, if derivatives generated by only numerical procedure is available then one should use a non-derivative Minimization procedure as it is more efficient [3].

Although, in some situations it is always preferable to use derivatives generated though numerical procedure. By following way these can be generated: suppose there is a small alteration (δx_i) in one of the coordinates x_i and the energy calculation is performed for this new alteration the by dividing the alteration in energy (δE) by the alteration in coordinate ($\delta E / \delta x_i$), the derivative $\partial E / \partial x_i$ is obtained. This rigorously yields the derivative at the mid-point between the two points x_i and $x_i + \delta x_i$. A more correct value of the derivative at the point x_i ; could also be acquired (at the price of a further energy calculation) by assessing the energy at two points, $x_i + \delta x_i$ and $x_i - \delta x_i$. The derivative is then obtained by dividing the variation with in the energies by $2\delta x_i$.

2. Non-derivative minimization methods

2.1 The simplex method

A geometrical figure having $M + 1$ interconnected vertices is called as *simplex*, where dimensionality of the energy function is M . Thus, a simplex with function of two variables will have a triangular shape. Further, for a function of three variables simplex will have tetrahedral shape. Therefore, for an energy function of $3N$ Cartesian coordinates the simplex will have $3N + 1$ vertices; but simplex will have $3N - 5$ vertices, if internal coordinates are used. The energy could be calculated for a specific set of coordinates correspond to each every vertex. For the function $f(x,y) = x^2 + 2y^2$ the simplex method would use a triangular simplex [5].

The simplex algorithm identifies an energy minimum by traveling around on the potential energy surface in a manner that is similar to the movement of an amoeba. There are three possible primary moves. The most common move is a reflection of the vertex having maximum value on the opposite sides of the simplex. The reflection is used as an effort to produce a new point having a lower value. If this is the lowest energy point than any other points in the simplex then next move may be applied which is a “reflection and expansion.” Reflection move will be failed to generate a better point, when a “floor of the valley” is reached. In this situation, simplex will make simple contraction all along the highest point dimension. If this fails to further decrease the energy then another kind of move is possible. In this move, contractions occur in all the directions towards the lowest point. The **Figure 2** illustrates above discussed three moves.

The vertices of the initial simplex have to be first generated before applying the simplex algorithm. The first conformation of the method fit to just one of these vertices. Rest of the points can be generated using various methodologies, e.g.

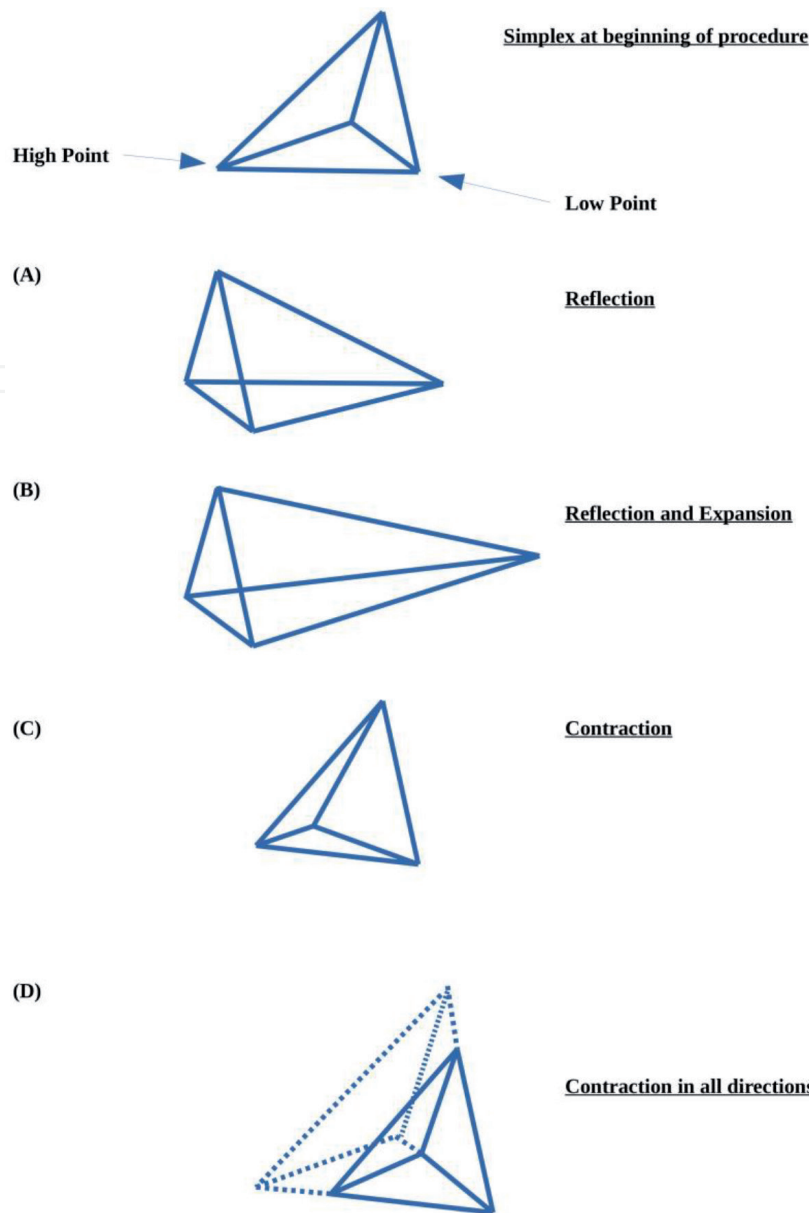


Figure 2.
The three basic moves permitted to the simplex algorithm (reflection, and its close relation reflect-and-expand; contract in one dimension and contract around the lowest point).

simplest method is to increase a fix value to each coordinate successively. To calculate the functional value of the applicable vertex, the energy of the whole system is measured for each new point.

When the starting configuration of the system is having high energy, it is best to use simplex method. The simplex method is more helpful in this because it seldom go wrong in the identification of a fitter answer. Nonetheless, it requires large computational time for the analysis of the high number of energy instances. For e.g. to create the starting simplex needs $3N + 1$ energy analysis. Due to this, the simplex method is frequently used along with other Minimization algorithms. In practice, starting configuration is fine tuned with few steps of the simplex method and then a more suitable and efficient method can be used for further calculations [6].

An important question is that what is the reason behind containing one extra vertex in the simplex than the degree of freedom? The answer to this is that of simplex is having lesser vertices than $M + 1$ then the simplex algorithm cannot search the entire surface of the energy. For e.g. if the simplex having just two vertex (a simplex with only two vertices is simply a straight line) is being used to search the quadratic surface of the energy, the only available move in this scenario would be

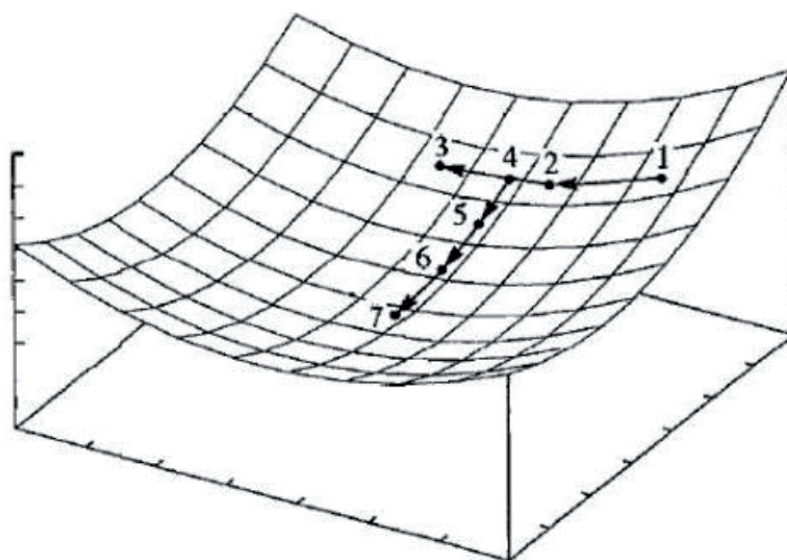


Figure 3.

*The sequential univariate search procedure. From the starting point 1, two steps are created along one of the coordinates to give points 2 and 3. A parabola is fitted to these three points and the minimum located (point 4). The same steps are then repeated along the next coordinate (points 5, 6 and 7) (Figure adapted from Schlegel H B 1987. Optimization of equilibrium geometries and transition structures In Lawley K P (editor) *ab initio methods in quantum chemistry* - I New York, John Wiley, pp. 249–286).*

to find out other points that lie on this straight line. In this case, the energy surface which is away from the straight line would not be searched. Likewise, if we have function of three variable and simplex is just a triangle then only the area of search space that lies in the same plane as to the triangle will only be searched, whereas the energy minimum may not be present at this plane [7].

2.2 The sequential univariate search method

It is seldom appropriate to use the simplex method for the calculations involve in quantum mechanics because in that case very high number of energy assessments has to be done. In this case a much befitting non-derivative procedure like the sequential univariate search method is well-advised [8]. This procedure consistently repeats through the coordinates successively. For every coordinate, two new configurations are created by making changes in the present coordinates (i.e. $x_j + \delta x_i$, and $x_i + 2\delta x_i$). Then the energy calculation for these two configurations is performed. Three points related to the two twisted configurations and the original one are then fitted with a parabola. The identification of the minimum point in the current quadratic function is performed. Then in the next step, the coordinate is twisted to the point of the minimum. The procedure is illustrated in **Figure 3**.

The minimum is bound to reach when the changes in all the coordinates are adequately very small. Alternatively, a new iteration is performed. In comparison to the simplex method, the sequential univariate method normally needs less function assessment. But if two or more coordinates have a strong connection or bonding then the sequential univariate search method may converge slowly. It also converges slowly when the energy surface is similar to a long narrow valley.

3. Derivative minimization methods

Most of the favorite Minimization procedures utilize derivatives because the information which is helpful in minimization is furnished by derivatives. The direction of the first derivative of the energy (the gradient) points where the minimum

lies and the magnitude of the gradient tells about the steepness of the local slope. The energy of the system can be decreased by moving each atom with respect to the force acting on it. The force is equal to minus the gradient. Second derivatives point towards the curvature of the function. This information can be utilized to find out where the function will change its direction (i.e. pass through a minimum or any different non-moving point).

The energy functions which are often utilized in molecular modeling are seldom quadratic and thus the Taylor series expansion can only be a well-advised approximation. There are two crucial consequences of this. First is, for a pure quadratic function a given minimization procedure executes very well rather than for a molecular mechanics or quantum mechanics energy surface. For example, the Newton–Raphson algorithm can identify the minimum in a one step for a purely quadratic function. But, for a typical molecular modeling energy function, it needs to run several iterations. The second consequence is that, even though they may function very well close to a minimum, where the harmonic approximation is more logical, the harmonic approximation is very bad and far from minimum. Due to this some of the less robust methods will not be successful. Because of this reason Minimization protocol must be picked very carefully. A robust or may be inefficient method could be exploited earlier than a comparatively least robust but more efficient procedure.

On the basis of highest order derivatives used, the derivative methods can be classified. The first derivatives or the gradients based methods are called as first-order methods. Methods in which both first and second order derivatives are used are known as second-order methods. Because the simplex method does not use any derivatives can thus be called as a zeroth-order method.

3.1 First-order minimization methods

The *steepest descents* and the *conjugate gradient* method are two such first order Minimization algorithms which are very often used in molecular modeling. In these methods coordinates of the atoms are altered step by step with respect to their movement towards the minimum point. For each iteration (k), the initial point is the molecular conformation generated from last step. It is represented by the multidimensional vector $x_k - 1$. For the first iteration, the starting point is the initial configuration of the system provided by the user, the vector x_1 .

3.1.1 The steepest descents method

The steepest descents method moves in the direction parallel to the net force, which in our geographical analogy corresponds to walking straight downhill. For $3N$ Cartesian coordinates this direction is most conveniently represented by a $3N$ -dimensional unit vector, s_k . Thus:

$$s_k = -g_k / |g_k| \quad (2)$$

Once the direction of movement is clearly characterized then it should be decided that how much distance to be covered along the gradient. Consider the two-dimensional energy surface of **Figure 4**. The gradient direction from the initial point is along the line shown. Suppose we have a cross-section through the surface along the line, the function will pass through a minimum and then increase, as shown in the figure [9]. We can identify the minimum point by performing a line search or we can take a step of arbitrary size along the direction of the force.

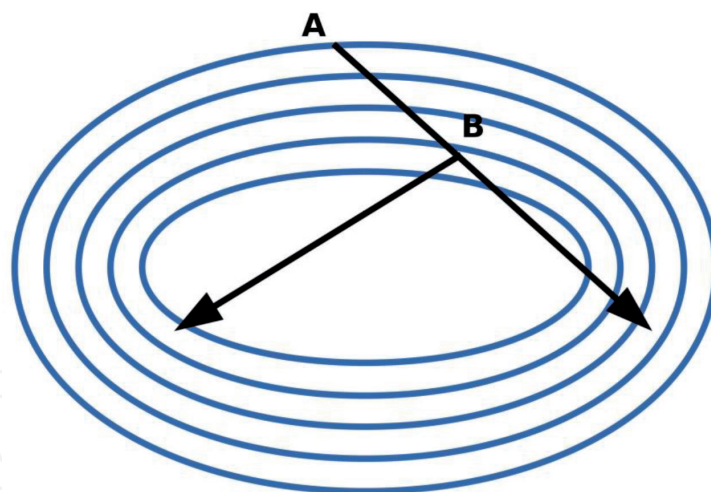


Figure 4.
Steepest descents method.

3.1.2 Line search in one dimension

The goal of a line search is to find out the minimum along a specific direction (i.e. along a line through the multidimensional space) [10]. In the very first step of the line search is to *bracket* the minimum. This implies determining three points along the line in a way that the energy of the intermediate point is less than the energy of the two extrinsic points. If it is possible to identify these kinds of three points, then it should be make sure that two extrinsic points must have at least one minimum in between. Then to reduce the distance in between the three points, an iterative algorithm could be applied which in a step by step manner, limits the minimum to a very smaller space. Theoretically, it looks easy but it may involve a large number of functional analysis. Thus it is computationally very expensive methods.

Alternatively, we can set a suitable quadratic function to the three points. Then apply differentiation to this suited function to modify an approximation to the minimum along the line which should be identified analytically. To get a better approximate, a new function can be set then, as shown in **Figure 5**. Higher-order polynomials may yield an improved fit to the bracketing points but when these

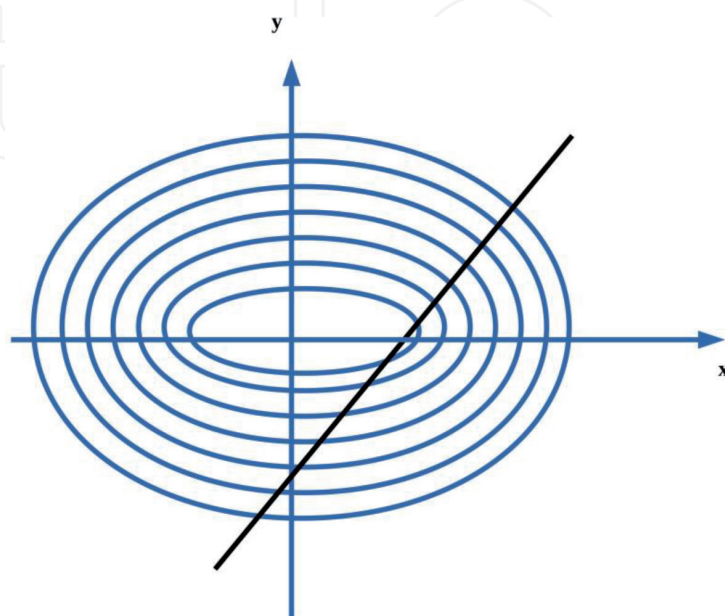


Figure 5.
A line search is used to locate the minimum in the function in the direction of the gradient.

are utilized with functions that altered aggressively in the bracketed region, these higher-order polynomials can yield wrong interpolations. The gradient at the minimum point obtained from the line search will be perpendicular to the previous direction. Thus, when the line search method is used to locate the minimum along the gradient then the next direction in the steepest descents algorithm will be orthogonal to the previous direction (i.e. $g_k \cdot g_{k-1} = 0$) [11].

3.1.3 Arbitrary step approach

As, we know that the line search may be computationally very expensive, New coordinates can be identified by walking a step of arbitrary size along the gradient unit vector s_k . The new set of coordinates after step k would then be given by the equation:

$$x_{k+1} = x_k + \lambda_k s_k \quad (3)$$

where, λ_k is the step size. In most of the applications within molecular modeling, the steepest descents algorithm, the step size at the start has a predetermined default value. If energy decreases after the first iteration, then for second iteration the step size is increases by an increasing component. The process repeats till the point at which each iteration decreases the energy. When a step produces an addition in energy, it is assumed that the algorithm has leapt across the valley which comprise the minimum and up the slope on the opposite face. The step size is then reduced by a multiplicative factor (e.g. 0.5). Often, the size of the step is decided according to the nature of the energy surface. It would be more suitable to have bigger step size for a plane or flat surface rather than a slender or narrow altered valley, where more smaller step are much appropriate. Computational time is less in the case of the arbitrary step method than much stringent line search method, because the arbitrary or random step approach needs higher number of steps to find out the minimum than line search method but arbitrary step method may frequently needs lesser functional analysis [12].

The largest inter-atomic forces indicate the direction of the gradient. Therefore, the steepest descent is more suitable for alleviating attributes of the highest-energy in the initial conformation. If the harmonic calculations corresponding to the energy is hypothesized badly and the initial point is distant from a minimum, even then the method performs strongly. But, in the case of downward movements in a long slender valley, the method uses short steps in high number and this causes trouble to the method. Although, it is not suitable manner to find out the minimum, the steepest descents process is bound to move in the right-angled direction at every point. The route constantly over compensates itself and vibrates. However, Subsequent steps reintroduce errors which were already rectified by prior steps [13].

3.1.4 Conjugate gradients minimization

The vibrating activity of the steepest descents procedure in slender depression is absent in the set of directions generated by the conjugate gradients methods. Rather, both the directions of consecutive steps and the gradients are orthogonal in the steepest descents method [14]. More specifically, in the conjugate gradients method, the gradients are orthogonal in nature at every point and the directions of consecutive steps are *conjugate* that is why it is more correctly known as the conjugate direction method. Because of the feature of a set of conjugate directions, for a quadratic function of M variables, in M steps the minimum can be identified. The conjugate gradients method moves in a direction v_k from point x_k where v_k is computed from the gradient at the point and the previous direction vector v_{k-1} .

Both the conjugate gradients method and the steepest descents method move in the direction of the gradient in the first step. The line search method should ideally be used to find out the one-dimensional minimum in all direction to assure that each gradient is orthogonal to all preceding gradients and that each direction is conjugate to all preceding directions. However, at this stage random step procedure is also achievable [15]. To identify the second point a line search should be applied along the line with gradient but it must pass through the point. Therefore, the conjugate gradient procedure identifies the perfect minimum of the function in just two moves.

3.2 Second order derivative methods

3.2.1 The Newton-Raphson method

Second-order methods utilize the information from both the first derivatives and the second derivatives to find out a minimum. First derivatives provide gradient information while second derivative furnish details about the curvature of the function. Having these properties, the Newton–Raphson method is the simplest second order method [16]. For a strictly quadratic function of the first derivative the second derivative will be same everywhere. If we talk about a multidimensional function the Hessian matrix of second derivatives essentially be inverted. Thus, for larger molecules it is more computationally expensive because there are a large number of atoms present and this necessitates bigger storage. The Newton–Raphson method is thus more appropriate to small molecules (usually less than 100 atoms or so) [17].

As stated earlier, for a strictly quadratic function, the Newton–Raphson method requires just one step to locate the minimum from any point on the surface. Practically, the surface is exclusively quadratic to a first approximation and this necessitates a large number of steps to move. The Hessian matrix of second derivatives should be calculated first and then inverted at each step. This must be ‘positive definite’ in a Newton–Raphson Minimization method. A positive definite matrix is one for which all the eigen-values are positive. When the Hessian matrix is not positive definite then the Newton–Raphson method moves to saddle points where the energy increases, rather than a narrow point where energy decreases. Additionally, the harmonic approximation is not suitable at positions which are very far from the minimum because this leads to instability of the Minimization. This can be solved by employing a more efficient and robust method (prior to the application of the Newton–Raphson method) to find out minimum or to reach close to minimum (in case of the positive definite Hessian matrix) [18].

3.2.2 Quasi-Newton method

Computation of the inverse Hessian matrix can be a possibly long procedure that represents an important disadvantage to the ‘pure’ second derivative methods such as Newton–Raphson. Furthermore, analytical second derivatives could not be generated preferably. Variable metric methods which are also an alternative name to the Quasi-Newton methods gradually develop the inverse Hessian matrix in consecutive iterations. That means, a sequence of matrices H_k is developed.

At each iteration k , the new positions x_{k+1} are obtained from the current positions x_k , the gradient g_k and the current approximation to the inverse Hessian matrix H_k . For quadratic function it is same, but for ‘real’ job a line search may be desired. Hence, a line search is performed along the vector $(x_{k+1} - x_k)$. It may not be essential to find out the minimum in the direction of the line search very accurately, at the cost of a few more steps of the quasi-Newton algorithm [19]. For quantum

mechanics calculations the additional energy evaluations required by the line search may prove more expensive than using the more approximate approach. An effective compromise is to fit a function to the energy and gradient at the current point x_k and at the point x_{k+1} and find out the minimum in the fitted function [20].

4. Which minimization method should I use?

The choice of Minimization algorithm is determined by a number of components, including the storage and computational requirements, the relative speeds with which the various parts of the calculation can be performed, the availability of analytical derivatives and the robustness of the method. Thus, any method that requires the Hessian matrix to be stored (let alone its inverse calculated) may present memory problems when applied to systems containing thousands of atoms. Calculations on systems of this size are invariably performed using molecular mechanics, and so the steepest descents and the conjugate gradients methods are very popular here. For molecular mechanics calculations on small molecules, the Newton–Raphson method may be used, although this algorithm can have problems with structures that are far from a minimum. For this reason it is usual to perform a few steps of Minimization using a more robust method such as the simplex or steepest descents before applying the Newton–Raphson algorithm. Analytical expressions for both first and second derivatives are available for most of the terms found in common force fields. The steepest descent method can actually be superior to conjugate gradients when the starting structure is some way from the minimum. However, conjugate gradients are much better once the initial strain has been removed. Quantum mechanical calculations are restricted to systems with relatively small numbers of atoms, and so storing the Hessian matrix is not a problem. As the energy calculation is often the most time-consuming part of the calculation, it is desirable that the Minimization method chosen takes as few steps as possible to reach the minimum. For many levels of quantum mechanics theory analytical first derivatives are available. However, analytical second derivatives are only available for a few levels of theory and can be expensive to compute. The quasi-Newton methods are thus particularly popular for quantum mechanical calculations.

When using internal coordinates in a quantum mechanical Minimization it can be important to use an appropriate Z-matrix as input. For many systems the Z-matrix can often be written in many different ways as there are many combinations of internal coordinates. There should be no strong coupling between the coordinates. Dummy atoms can often help in the construction of an appropriate Z-matrix. A dummy atom is used solely to define the geometry and has no nuclear charge and no basis functions. Strong coupling between coordinates can give long ‘valleys’ in the energy surface, which may also present problems. Care must be taken when defining the Z-matrix for cyclic systems in particular. The natural way to define a cyclic compound would be to number the atoms sequentially around the ring. However, this would then mean that the ring closure bond will be very strongly coupled to all of the other bonds, angles and torsion angles. Some quantum mechanics programs are able to convert the input coordinates (be they Cartesian or internal) into the most efficient set for Minimization so removing from the user the problems of trying to decide what is an appropriate set of internal coordinates. For energy Minimizations redundant internal coordinates have been shown to give significant improvements in efficiency compared with Cartesian coordinates or non-redundant internal coordinates, especially for flexible and polycyclic systems [21]. The redundant internal coordinates employed generally comprise the bond lengths, angles and torsion angles in the system. These methods obviously also

require the means to inter-convert between the internal coordinate representation and the Cartesian coordinates that are often used as input and desired as output. Of particular importance is the need to transform energy derivatives and the Hessian matrices (if appropriate) [22].

5. Differentiating between minima, maxima and saddle points

A configuration at which all the first derivatives are zero need not necessarily be a minimum point; this condition holds at both maxima and saddle points as well. From simple calculus we know that the second derivative of a function of one variable, $f''(x)$ is positive at a mini- minimum and negative at a maximum. It is necessary to calculate the eigenvalues of the Hessian matrix to distinguish between minima, maxima and saddle points. At a minimum point there will be six zero and $3N - 6$ positive eigenvalues if $3N$ Cartesian coordinates are used. The six zero eigenvalues correspond to the translational and rotational degrees of free- freedom of the molecule (thus these six zero eigenvalues are not obtained when internal coordinates are used). At a maximum point all eigenvalues are negative and at a saddle point one or more eigenvalues are negative.

6. What should be the convergence criteria?

In contrast to the simple analytical functions that we have used to illustrate the operation of the various Minimization methods, in 'real' molecular modeling applications it is rarely possible to identify the 'exact' location of minima and saddle points. We can only ever hope to find an approximation to the true minimum or saddle point. Unless instructed otherwise, most Minimization methods would keep going forever, moving ever closer to the minimum. It is therefore necessary to have some means to decide when the Minimization calculation is sufficiently close to the minimum and so can be terminated. Any calculation is of course limited by the precision with which numbers can be stored on the computer, but in most instances it is usual to stop well before this limit is reached. A simple strategy is to monitor the energy from one iteration to the next and to stop when the difference in energy between successive steps falls below a specified threshold. An alternative is to monitor the change in coordinates and to stop when the difference between successive configurations is sufficiently small. A third method is to calculate the root-mean-square gradient. This is obtained by adding the squares of the gradients of the energy with respect to the coordinates, dividing by the number of coordinates and taking the square root. It is also useful to monitor the maximum value of the gradient to ensure that the Minimization has properly relaxed all the degrees of freedom and has not left a large amount of strain in one or two coordinates [23].

7. Applications of energy minimization

Energy Minimization is very widely used in molecular modeling and is an integral part of techniques such as conformational search procedures. Energy Minimization is also used to prepare a system for other types of calculation. For example, energy mini- Minimization may be used prior to a molecular dynamics or Monte Carlo simulation in order to relieve any unfavorable interactions in the initial configuration of the system [24]. This is especially recommended for simulations of complex systems such as macromolecules or large molecular assemblies.

8. Conclusion

The energetic state of a protein is one of the most important representative parameter of its stability. The energy of a protein (E) can be defined as a function of its atomic coordinates, thus providing a quantitative criterion for model selection and refinement. This energy function consists of several components e.g. (1) Bond energy and angle energy, representative of the covalent bonds, bond angles. (2) Dihedral energy, due to the dihedral angles. (3) A van der Waals term (also called Leonard-Jones potential) to ensure that atoms do not have steric clashes. (4) Electrostatic energy accounting for the Coulomb's Law in protein structure, i.e. the long-range forces between charged and partially charged atoms. All these quantitative terms have been parameterized and are collectively referred to as the 'forcefield'. The goal of energy Minimization is to find a set of coordinates representing the minimum energy conformation for the given structure. Various algorithms have been formulated by varying the use of derivatives. The common algorithm used for this optimization is steepest descent, conjugate gradient and Newton-Raphson etc. These methods complement each other in search of the local minima. Therefore, a reasonable energy Minimization protocol involves few initial steps of steepest descent, followed by a larger number of conjugate gradient iterations. Although energy Minimization is a tool to achieve the nearest local minima, it is also an indispensable tool in correcting structural anomalies, viz. bad stereo-chemistry and short contacts. An efficient optimization protocol could be devised from these methods in conjunction with a larger space exploration algorithm, e.g. molecular dynamics.

Acknowledgements

The authors are grateful to the Sam Higginbottom University of Agriculture, Technology and Sciences, Allahabad, for providing the facilities and support to complete the present work.

Conflict of interest


The author declares no conflict of interest.

Author details

Budhayash Gautam
Sam Higginbottom University of Agriculture, Technology and Sciences,
Prayagraj, U. P., India

*Address all correspondence to: budhayashgautam@gmail.com

IntechOpen

© 2020 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Ebejer Jean-Paul, Fulle Simone, Morris Garrett M., Finn Paul W. The emerging role of cloud computing in molecular modelling, *Journal of Molecular Graphics and Modelling*, 2013; 44,177-187.
- [2] Adcock S A, McCammon J A. *Molecular dynamics: survey of methods for simulating the activity of proteins. Chem Rev.* 2006;106:1589-1615.
- [3] Hinchliffe Alan. *Molecular Modelling for Beginners*, 2nd Edition. John Wiley & Sons Ltd. 2003.
- [4] Leach AR. *Molecular Modelling: Principles and Applications*. Prentice Hall, 2001.
- [5] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, Clifford Stein. *Introduction to Algorithms*, Second Edition. MIT Press and McGraw-Hill, 2001;790-804.
- [6] Disser, Yann, Skutella Martin. The Simplex Algorithm Is NP-Mighty. *ACM Trans. Algorithms*. 2018;15(1):5:1-5:19.
- [7] Maros István, Mitra Gautam. *Simplex algorithms. In J. E. Beasley (ed.). Advances in linear and integer programming. Oxford Science.* 1996;1-46.
- [8] Soumitra Sitole. *Univariate Search Method (Optimizing Quadratic Equations with Two Variables)*, MATLAB Central File Exchange. 2020.
- [9] Wiberg K J. A Scheme for Strain Energy Minimization. Application to the Cycloalkanes. *Am. Chem. Soc.*, 1965; 87,5,1070-78.
- [10] Nocedal Jorge, Wright Stephen J. *Line Search Methods. Numerical Optimization*. New York: Springer. 1999; 34-63.
- [11] Wenyu Sun, Ya-Xiang Yuan. *Line Search. Optimization Theory and Methods: Nonlinear Programming*. New York: Springer. 2006;71-117.
- [12] Elber R, Meller J , Olender R. Stochastic Path Approach to Compute Atomically Detailed Trajectories: Application to the Folding of C Peptide *The Journal of Physical Chemistry B.* 1999;103(6),899-911.
- [13] Hsieh W, Kuo M, Yau H F, Chang Chi Ching. A simple arbitrary phase-step digital holographic reconstruction approach without blurring using two holograms. *OPT REV.* 2009;16,466-471.
- [14] Fletcher R, Reeves C M. Function Minimization by conjugate gradients. *The Computer Journal.* 1964;7,2,149-154.
- [15] Polak E, Ribiere G. *Rev. Fr. Infolrm. Rech. Operations* 16R1,1969;35-43.
- [16] Ypma, Tjalling J. Historical development of the Newton–Raphson method, *SIAM Review.* 1995;37(4), 531-551.
- [17] Gil A, Segura J, Temme N M. *Numerical methods for special functions. Society for Industrial and Applied Mathematics (SIAM).* 2007; <http://dx.doi.org/10.1137/1.9780898717822>
- [18] Süli Endre, Mayers David. *An Introduction to Numerical Analysis. Cambridge University Press*, 2003.
- [19] Broyden C G. *Quasi-Newton Methods. In Murray, W. (ed.). Numerical Methods for Unconstrained Optimization. London: Academic Press.* 1972;pp.87-106.
- [20] Haelterman Rob, Eester Dirk Van, Verleyen Daan. Accelerating the solution of a physics model inside a tokamak using the (Inverse) Column Updating Method. *Journal of Computational and Applied Mathematics.* 2015;279:133-144.

[21] Peng C, Ayala PY, Schlegel H B, Frisch M J. Using Redundant Internal Coordinates to Optimise Equilibrium Geometries and Transition States. *Journal of Computational Chemistry*. 1996;17:49-56.

[22] Westheimer F H. Steric Effects in Organic Chemistry, ed. M. S. Newman, John Wiley & Sons, New York, 1956.

[23] Hendrickson J B. Molecular Geometry. I. Machine Computation of the Common Rings. *J. Am. Chem. Soc.* 1961; 83, 22, 4537-47.

[24] Schlegel H B. Optimization of Equilibrium Geometries and Transition Structures In Lawley K P (Editor) *Ab Initio Methods in Quantum Chemistry - I* New York, John Wiley, 1987; 249-286.