Journal of Computer-Aided Molecular Design, 1 (1987) 197–210
ESCOM

J-CAMD 021

A method for fast energy estimation and visualization of protein-ligand interaction

Nobuo Tomioka, Akiko Itai and Yoichi Iitaka

Faculty of Pharmaceutical Sciences, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan 113

Received 12 August 1987
Revised 16 November 1987
Accepted 17 November 1987

Key words: Computer graphics; Docking simulation; Grid point; Empirical potential function

SUMMARY

A new computational and graphical method for facilitating ligand-protein docking studies is developed on a three-dimensional computer graphics display. Various physical and chemical properties inside the ligand binding pocket of a receptor protein, whose structure is elucidated by X-ray crystal analysis, are calculated on three-dimensional grid points and are stored in advance. By utilizing those tabulated data, it is possible to estimate the non-bonded and electrostatic interaction energy and the number of possible hydrogen bonds between protein and ligand molecules in real time during an interactive docking operation. The method also provides a comprehensive visualization of the local environment inside the binding pocket.

With this method, it becomes easier to find a roughly stable geometry of ligand molecules, and one can therefore make a rapid survey of the binding capability of many drug candidates. The method will be useful for drug design as well as for the examination of protein-ligand interactions.

INTRODUCTION

Recently, studies on protein chemistry and drug design using atomic coordinates of protein structures obtained from crystal structure analyses have become more and more popular, due to the remarkable developments in protein crystallography and in isolation techniques for proteins.

Results of structure analyses on the complexed state between proteins and ligands are especially useful to deepen our understanding about protein-ligand interactions and molecular recognition in the biological systems. Such protein structures are used to simulate enzyme reactions or interactions with natural substrates [1, 2]. By such docking simulations, not only differences in substrate specificities and/or rates of enzymatic reactions are explained, but also inhibitors with stronger affinity to the target enzyme can be designed [3, 4].

A ligand molecule which can bind strongly to the target protein should have energetically favorable interactions with the protein with an appropriate relative geometry. In the docking simulation, the problem of finding such geometry between two molecules is too difficult to be accom-

plished only by computational methods. Even without conformational freedom, six degrees of freedom for rotation and translation of ligand may give rise to innumerable local minima, from which a global minimum cannot be easily discriminated. Therefore, for the time being, likely stable geometies are usually selected by human judgment with a three-dimensional graphics display.

During the last ten years, many docking simulation studies using three-dimensional graphics have been published in order to find stable relative geometries for ligand molecules, starting from the known structure of a complex between a protein and another ligand molecule.

In 1981, a method for representing electrostatic potentials on molecular surfaces was presented by Weiner et al. [5]. They calculated electrostatic potential values on the solvent accessible surface from all partial atomic charges of each molecule, and displayed the potential value by color-coding the molecular surface according to the potential level. Using several examples of ligand-protein complexes whose structures had been elucidated by X-ray structure analyses, they showed good complementarity in shape and electrostatic properties between a ligand and its target enzyme. Moreover, they suggested that the method might be useful for docking studies for designing drugs [6].

However, their method gives no numerical index to indicate the goodness of fit between ligand and protein. Although the protein-ligand interaction energy is a good indicator in selecting or modeling ligand molecules with stronger affinity to the target protein, energy calculations by the conventional atom-pair method are not fast enough for interactive docking simulations. In addition, the conformational freedom of non-rigid ligand molecules makes the problem much more difficult and time-consuming. Therefore, a new efficient method of estimating the interaction energy between two molecules is required for interactive docking studies.

We have developed a new program system GREEN for the purpose of high-speed and rational docking studies [7,8]. We use precalculated grid point data which holds physical and chemical properties inside the ligand binding cavity of protein molecules. In 1985, Goodford has presented a similar approach which also uses grid point data for representations of favorable sites of ligand functional groups [9]. In our method, grid point data are used not only for representations but also for estimation of ligand-protein interaction energy.

Our program system enables a real-time estimation of interaction energies between a ligand and a protein, and also offers comprehensive representations of the environment around the ligand binding site during the interactive fitting operation on a three-dimensional graphics display. The representations facilitate an initial introduction of a new ligand molecule into the binding site, and the rapid energy estimation accelerates the fitting operation and quickly leads to a reasonable result. The roughly determined ligand geometry thus obtained can be stored in a file and, if necessary, refined by molecular mechanics calculation using programs for macromolecular systems such as AMBER [10] or CHARMM [11].

MATERIALS AND METHODS

Preparation of data

Atomic coordinates of proteins are taken from the Protein Data Bank. An appropriate choice is made where alternative atomic positions have been provided. A cofactor bound in a complexed crystal is treated as a part of the protein if it is structurally and functionally important. Some of

the water molecules or ions around the ligand binding pocket are also included as protein atoms after a careful examination of hydrogen bonding patterns and the crystallographic B-factor.

Although hydrogen positions are missing from most of the Protein Data Bank entries, they are important for assessing the accurate size and shape of ligand binding pockets. Moreover, hydrogen atoms attached to heteroatoms bear significant positive atomic charges and therefore should not be ignored for the proper electrostatic treatments of proteins. We compute all hydrogen atom positions around the binding pocket according to standard geometry. Orientation of hydrogen atoms attached to rotatable functional groups, such as hydroxyl, are adjusted on the consideration of hydrogen bonds to other functional groups.

An atomic charge is automatically allotted to every protein atom using the standard value for amino acid atom. We use the values published by Weiner et al. [12].

Atomic coordinates of ligand molecules are either taken from the Cambridge Crystallographic Database, or are generated by an interactive molecular modeling program (unpublished) on a three-dimensional graphics display. Atomic charges on ligand atoms are assigned by molecular orbital calculation. As the charge distribution is not so much affected by changes of conformations of molecules, we use those initially calculated charge distribution throughout the docking study.

Calculation of grid point data

Three-dimensionally tabulated physical and chemical properties are used for visualization on a graphics display and/or real-time energy calculations.

A three-dimensional grid with a regular interval of 0.4–1.0 Å is generated inside the ligand binding pocket of a protein. Several types of probe atoms are placed on each grid point. Every atom species which might exist in the ligand molecules to be studied (e.g., carbon, hydrogen, nitrogen, oxygen) is adopted as a probe atom. Any grid point on which the probe atom has an unacceptably close contact to a protein atom is excluded from further calculation.

For each probe placement and for each probe type, the non-bonded interaction energy is calculated by conventional atom-pair calculations using an empirical potential function. We use a Lennard-Jones type function described as follows:

$$V_{NBD}{}^{i} = \sum_{j} (A / r_{ij}{}^{12} - B / r_{ij}{}^{6})$$

Here i denotes a serial number of the probe atom position and j denotes a serial number of the receptor atom. A and B are empirical parameters defining position and depth of potential minimum. We currently use parameters published by Weiner et al., which were reported to be well suited to the molecular mechanics calculations of macromolecular systems [12]. r_{ij} is the distance between the current probe position and the receptor atom j.

In addition to several arrays of non-bonded potential values for different probe atoms, an array of electrostatic potential values is calculated as follows:

$$V_{ELC}{}^{i} = \sum_{j} (q_j / \varepsilon r_{ij})$$

The notations i, j and r_{ij} are the same as in the non-bonded equation above. q_j is the charge on the receptor atom j, and ε is a dielectric constant. The value of $V_{ELC}{}^{i}$ is equivalent to the electro-

static interaction energy between a positive unit charge placed on the current probe position and all atomic charges on the receptor atoms.

The choice of dielectric constant models are important for the proper treatment of the electrostatic interaction. A constant value model is often used for its simplicity but is not very realistic. We use a distance-dependent approximation of the dielectric constant: i.e., $\varepsilon = 4r_{ij}$ [13]. The approximation is still insufficient because it ignores the effect of water surrounding the protein molecule, but is better than the constant dielectric model [14, 15]. Furthermore, the distance-dependent model has an advantage that the amount of calculation required is minimum among the various dielectric models.

Another array is prepared for expressing the hydrogen bonding properties. For each grid point, our program checks the possibility of hydrogen bonding from all the hydrogen bonding functional groups in the protein molecule.

On the basis of statistical studies of hydrogen bonds observed in the crystallographic data base [16, 17], we have developed the following geometrical criteria on hydrogen bonding possibilities. Between the current probe location and every hydrogen bonding atom of the protein, we consider the distance DP (D: hydrogen donor atom, P: grid point) and/or distance AP (A: hydrogen acceptor atom) and angle DHP and/or ALP (L: lone pair). The distance allowance for hydrogen bonding is tentatively defined as from 2.5 to 3.1 Å. Angles are evaluated as the deviation from linearity and permissive deviation for hydrogen bonding is taken as 30°.

When these criteria are satisfied, our program sets a hydrogen bonding flag at the grid point according to the hydrogen bonding properties of the functional group affecting the grid point; i.e. hydrogen donor, hydrogen acceptor or ambivalent.

Positions of lone-pairs are not present in the Protein Data Bank data. Our grid point data preparation program sets lone-pair directions geometrically for each hydrogen bonding functional group, such as carbonyl, hydroxyl, amino and so on. We have defined 19 hydrogen bonding group categories in Table 1, and the category numbers are assigned to each atom during the data preparation step for the protein.

Arrays of calculated grid point data, including non-bonded, electrostatic and hydrogen bonding terms, are stored into a disk file and are used repeatedly in later visualization and energy esti-

TABLE I
NINETEEN HYDROGEN BONDING GROUP CATEGORIES

Category no.	Hydrogen bonding group	Category no.	Hydrogen bonding group
1	N primary amine (sp2)	11	O ether
2	N primary amine (sp3)	12	O carbonyl
3	N ammonium ion (sp3)	13	O carboxylate anion
4	N amide (sp2)	14	O carboxylic acid
5	N secondary amine (sp3)	15	O phosphate
6	N aromatic	16	O water molecule
7	N aromatic protonated	17	S mercapto
8	N tertiary amine	18	S thioether
9	N tertiary amine protonated	19	O hydroxyl (hydrogen position fixed)
10	O hydroxyl (C-O bond rotation allowed)		

mation. Three-dimensional locations of grid points in those arrays are also stored in a packed form into another index array.

Visualization

Using the non-bonded interaction term of the grid point data, we define an 'atom acceptable region'. The region is defined as a group of grid points whose values of the non-bonded interaction are negative. As the non-bonded term is prepared for each probe atom type, the region can be defined for each probe atom type.

By three-dimensionally contouring the non-bonded term V_{NBD} at a zero potential level, we get a bird-cage expression of the 'atom acceptable region'. According to the electrostatic interaction term V_{ELC} of the grid point data, different levels of color-coding can also be given to the bird-cage expression. The 'atom acceptable region' is displayed on the three-dimensional graphics and facilitates manual docking operation of ligand molecules to the receptor.

We also define a 'hydrogen bonding region' using the hydrogen bonding term of the grid point data. It is defined as a group of grid points whose hydrogen bonding flags are set. If a hydrogen bonding partner exists in this region, then a strong hydrogen bonding interaction would be expected between the partner and the protein.

The 'hydrogen bonding region' is displayed also as a bird-cage expression, surrounding the grid points on which hydrogen bonding flags are set. This cage is colored according to the type of the hydrogen bonding flag (donor, acceptor or ambivalent). Such a representation may help one to place hydrogen bonding functional groups when one is designing ligands with a specific hydrogen bonding capability.

Estimation of interaction energy

Using the three-dimensionally tabulated potentials and attributes calculated as described above, we can estimate the interaction energy between a ligand and the protein rapidly as follows.

The non-bonded interaction energy is estimated by summing up the non-bonded interaction terms on the grid point nearest to each ligand atom. According to the atom species of each ligand atom, the array for the corresponding probe atom is used for the calculation. Summation of the product of the atomic charge on each ligand atom and the electrostatic term V_{elc} on the nearest grid point gives an estimate of the electrostatic interaction energy.

The number of hydrogen bonds expected between the two molecules is counted by matching the type of each functional group in the ligand with the hydrogen bonding flag of the nearest grid point.

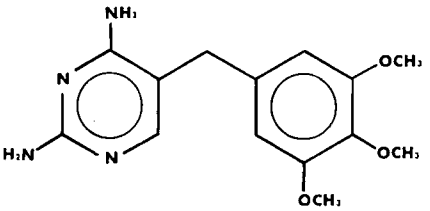


Fig. 1. Structure of trimethoprim.

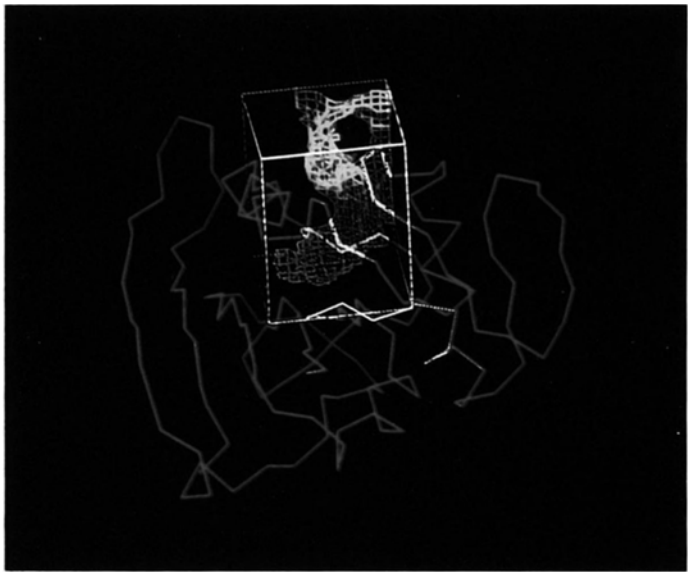


Fig. 2. Carbon 'atom acceptable region' in the substrate binding pocket of *E. coli* dihydrofolate reductase contoured at $V_{vdw} = 0$ kcal/mol. Colors of the cage indicate the electrostatic potential V_{elc} divided into nine levels with an interval of 12 kcal/mol. They are from blue (most negative; $V_{elc} < -42$ kcal/mol) to red (most positive; $V_{elc} > 42$ kcal/mol). The white rectangular box indicates a region of the grid point calculation. An α-carbon skeleton of the enzyme (purple) is also shown.

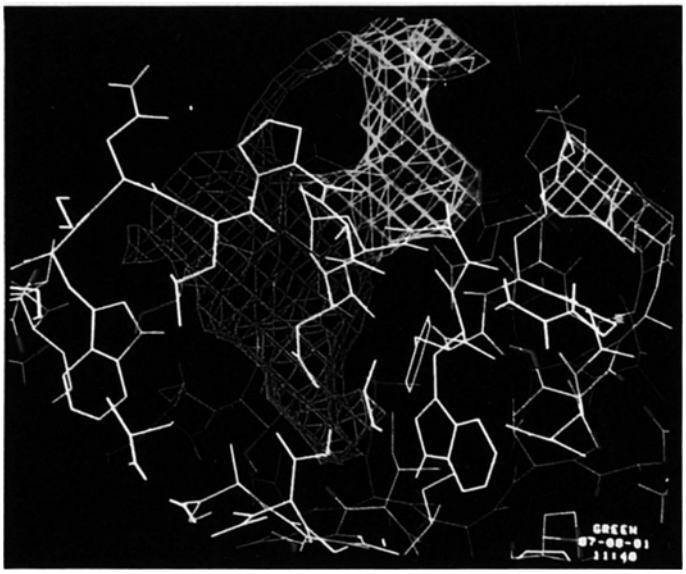
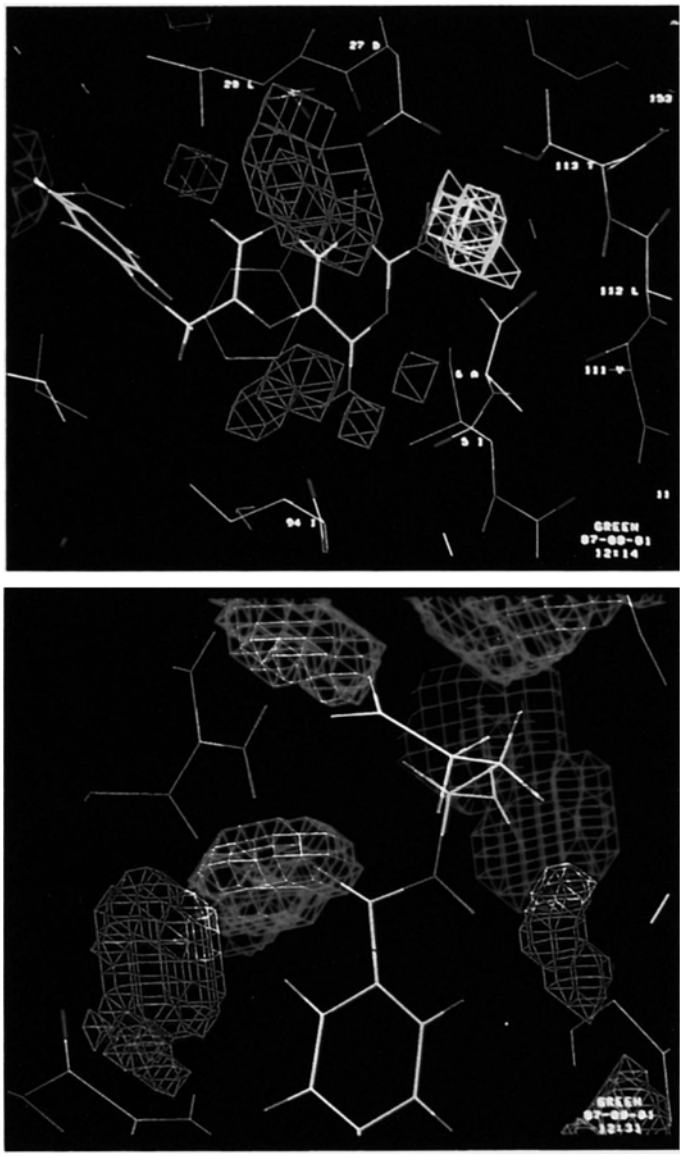


Fig. 3. 'Atom acceptable region' with a skeleton of *E. coli* dihydrofolate reductase. Conditions of drawing the 'atom acceptable region' are same as in Fig. 2. Color-coding of atoms are white: carbon, blue: nitrogen, red: oxygen, pale blue: hydrogen, purple: sulfur.

During the interactive docking operation, current ligand coordinates are sampled with short interval time (< 1 s), and interaction energy is calculated every time by the grid method. The total energy, together with the energy values for non-bonded interaction and electrostatic interaction and the number of possible hydrogen bonds are displayed in the upper part of the display screen.



a

b

Fig. 4. 'Hydrogen bonding region' of the substrate binding pocket of *E. coli* dihydrofolate reductase. Cages indicate the favorable locations of hydrogen bonding atoms of a ligand molecule which result from functional groups in the protein molecule. Cage colors indicate the type of the hydrogen bonding functional group affecting the 'hydrogen bonding region'. Red is for a hydrogen donor such as amide nitrogen, blue is for a hydrogen acceptor such as carbonyl oxygen, and yellow is for ambivalent functional groups such as hydroxyl oxygen. Enzyme and bound methotrexate (carbon in yellow) skeletons are also shown. a: Around the pteridine ring of methotrexate. b: Around the glutamate moiety of methotrexate.

Finding energy minimum

During the docking operation, the estimated energy values are not only displayed on the screen but also stored into a file. Rotation, translation and bond rotation angles of ligand molecule are also stored. With this 'logging' file, our program finds minimum energy geometry of the ligand molecule whenever we want.

The ligand geometry thus found is near but not exactly at the energy minimum. In order to

refine the ligand geometry, our program also incorporates simple steepest descent energy minimization. In the minimization, only intermolecular non-bonded and electrostatic energies are calculated by conventional atom-pair type method and the ligand molecule is moved as a rigid body.

More precise energy minimization should be done by external molecular mechanics programs such as AMBER or CHARMM.

Programs

Program system GREEN consists of three main programs (PDBFIL, CALGRD, MMGRD) and several smaller utility programs. PDBFIL is a protein data preparation program. Functions of PDBFIL are explained in the following section. Program for grid point calculation (CALGRD) is written for both VAX/VMS and HITAC supercomputer S-810 (highly vectorized version). Interactive program for docking simulation (MMGRD) was first developed for VAX/VMS and Evans & Sutherland PS-330 system and currently works on HITAC M-680H computer and COMTEC DS-351B graphics display. In addition to above two versions, we are also preparing an IRIS graphics workstation version of MMGRD. All programs are written according to the FORTRAN77 standard. (All programs are available on request.)

RESULTS

Dihydrofolate reductase

To illustrate our method, we have examined the interaction between dihydrofolate reductase and its inhibitor. Dihydrofolate reductase reduces substrate dihydrofolate to tetrahydrofolate using NADPH as a cofactor, and participates in the synthesis of thymidine. This enzyme is essential for all living things but its structure varies depending on species. Structures of three analogous enzymes of different origin have been elucidated by X-ray crystal analyses and have been compared with each other. Because of their great discrepancy in sensitivity to inhibitory drugs, many attempts have been made to design good inhibitors of these enzymes [18–20]. Some of the inhibitors have been used as anticancer or antibacterial drugs. One of them, trimethoprim (Fig. 1), is in clinical use. We have attempted to fit trimethoprim to the dihydrofolate reductase of *E. coli* for which the crystal structure analysis has been done as a binary complex with the inhibitor methotrexate [21, 22].

Atomic coordinates of *E. coli* dihydrofolate reductase were taken from the Protein Data Bank entry 4DFR. Atom types for non-bonded interaction energy function, atomic charges, hydrogen bonding functional group types and a bond table were automatically assigned to all the enzyme atoms by the program PDBFIL. Atomic charges of amino acids were taken from ab initio calculation results by Weiner et al. [12]. Hydrogen atoms were generated also by PDBFIL at geometrically appropriate positions. Some water molecules around the ligand binding site were included as enzyme atoms, judging their stability by their hydrogen bonding network and their crystallographic B-factor.

The coordinates of trimethoprim were taken from the Cambridge Crystallographic Data Base entry AMXBPM10 [23]. Atomic charges of trimethoprim were calculated by the ab initio MO calculation program GSCF2 [24] using STO-3G basis set.

Grid point data were calculated with a 0.4 Å grid interval within a $12.8 \times 16.8 \times 11.6$ Å^3 box-region including a substrate binding site of dihydrofolate reductase. Out of 42 570 possible grid

points, 11 965 grid points were selected and others were rejected because of severe close contacts between the probe atom and enzyme atoms. The probes used were hydrogen, carbon, nitrogen and oxygen atoms which exist in the trimethoprim molecule.

In Figs. 2 and 3, the 'atom acceptable region' for the carbon probe is displayed as a bird-cage expression. The cages are color-coded by electrostatic potential V_{elc} stored as grid point data. The expression is well suited to the fitting of a ligand molecule displayed as a wire frame model, since it indicates a region where the center of an atom can stably exist.

Fig. 4 shows a bird-cage expression of the 'hydrogen bonding region'. The expression indicates a region where functional groups of ligand molecules can make stable hydrogen bonds to enzyme atoms. This representation helps one to place hydrogen bonding functional groups of ligand molecules at adequate positions.

In Figs. 5 and 6, the trimethoprim molecule is fitted into the carbon 'atom acceptable region', by applying translation, rotation and bond rotation manually on the three-dimensional graphics display. The intermolecular interaction energy and the number of intermolecular hydrogen bonds are estimated by the grid method and are displayed on the screen while we are manipulating the trimethoprim molecule.

Calculation for the energy estimation is rapid enough for the estimated energy value to be updated at intervals of less than 1 s. As we manipulate the trimethoprim molecule on the graphics display, we can immediately judge the goodness of fitting by the energy estimation. The energy estimation speeds up the docking operation and leads it to a more reasonable result.

We have compared our docking result with the crystal structure of *E. coli* dihydrofolate reductase and trimethoprim binary complex[25]. Since the atomic coordinates of the binary complex is not available from the Protein Data Bank, the comparison was made by the protein-ligand close contact data and visual inspection of the stereo figure in the paper. In our model, the pyrimidine ring of trimethoprim is close to Ile-5, Ala-6, Ala-7, Asp-27, Leu-28, Phe-31, and Ile-94. The trimethoxybenzyl part is close to Leu-28, Phe-31, Ser-49, Ile-50, Arg-52, and Leu-54. Protein-ligand contacts of our model are almost consistent with the crystallographic result. Only one major exception is Arg-52, which is exposed outside of the ligand binding site and has different conformation between the methotrexate complex and the trimethoprim complex.

DISCUSSION

Docking simulation is one of the most practical uses of results from protein crystallography, and it is regarded as a typical means for the rational design of inhibitors or drugs. Developments of efficient methods for rapid docking simulation are desirable.

It is not easy to perform docking simulation between a protein and a ligand molecule only by a computational procedure. Even by using a three-dimensional graphics display, it needs many elaborate 'trials and errors' to find a stable geometry for the ligand. As well as the commonly used wire frame representation of a protein structure, a representation of spatial information such as the size, shape and local characteristics of the ligand binding site will greatly facilitate the docking study.

As a strategy for describing such spatial information inside the binding site, the three-dimensional grid point concept has often been used.

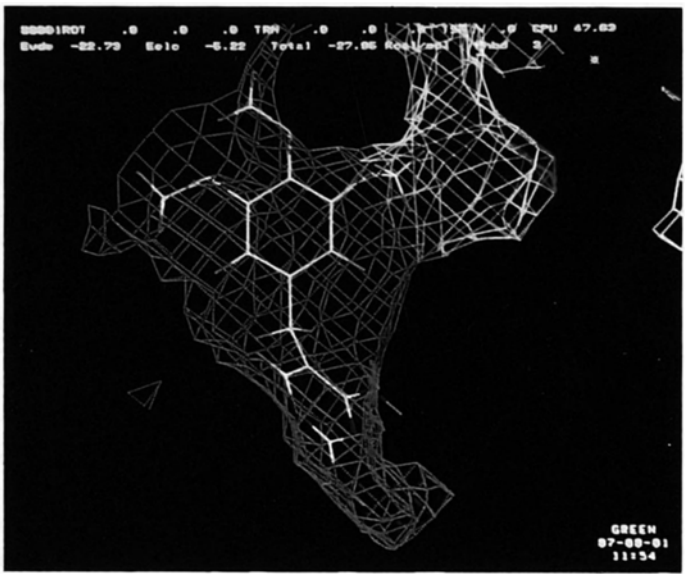


Fig. 5. Trimethoprim molecule on the 'atom acceptable region' of the carbon probe. The enzyme model is not displayed. The estimated values of the non-bonded and electrostatic interaction energies and the number of detected hydrogen bonds are displayed at the upper part of the screen. Those values are updated with an interval of less than 1 s while we are manipulating the trimethoprim molecule manually.

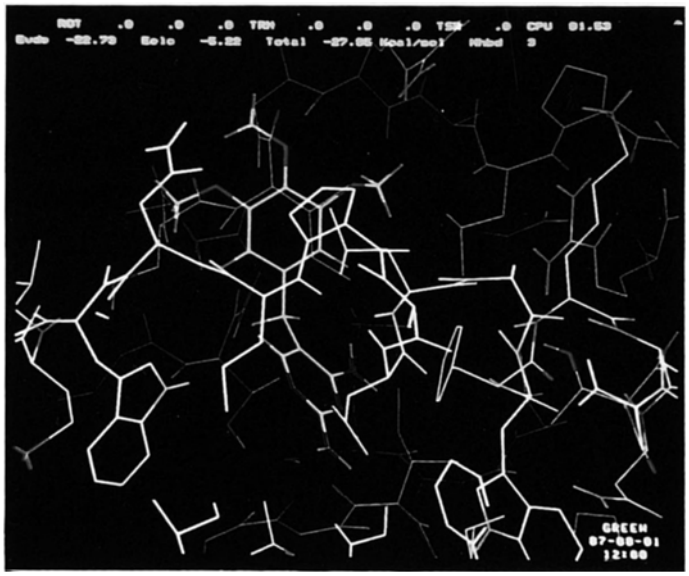


Fig. 6. The fitted trimethoprim molecule is shown with the skeleton of dihydrofolate reductase. Although the energy estimation can be performed with such a conventional representation of molecules, the 'atom acceptable region' representation is preferred in order to make the docking operation easy.

In 1983, Busetta et al. developed an interactive program for docking studies on a three-dimensional graphics display [26]. Their program indicated the allowed space for ligand molecules in order to avoid severe close contacts between ligand and protein. The allowed space was represented as the probe accessible volume which was calculated on a grid with regular interval inside

the ligand binding pocket. Although their program could perform non-interactive energy calculation or minimization at any stage of a docking process, it did not yield any energetic information in real time.

The three-dimensional tabulation of some kinds of potentials using systematic grid points is useful, not only for graphics visualization, but also for the estimation of intermolecular forces during interactive docking operations. Our method combines several advantageous features:

(1) By visualizing the environment of a ligand binding cavity of a protein, a new ligand molecule may be introduced and correctly placed into the binding cavity much more easily.

(2) The visualization enables the rational design of a ligand molecule which can bind more strongly and more specifically to the target protein, especially by examining hydrogen bonding between two molecules.

(3) By showing the intermolecular interaction energy between two molecules in real-time during the docking process, stable geometries of ligands are found more easily.

With such advantages, our method allows one to roughly examine many drug-candidate molecules in a short period, in order to find whether those molecules have possibilities of stable binding to the target protein.

Goodford has succeeded in showing favored sites of several functional groups (such as amine, hydroxyl, methyl, water and so on) inside a binding pocket of a protein [9]. Assuming each functional group to be a 'united' probe atom, the interaction energy with the protein was calculated on every three-dimensional grid point. Van der Waals interactions, electrostatic interactions and hydrogen bonding interactions were included into the total interaction energy between the probe functional group and the protein atoms. Non-conventional and elaborate considerations were paid in order to assess the electrostatic interaction. With regard to hydrogen bonding, Goodford used a direction dependent 6-4 potential which takes into account an angle DHP (D: donor heteroatom, H: hydrogen atom, P: probe atom) and a distance DP. The favorable sites were visualized by a bird-cage expression, which was obtained by contouring the energy map at a certain energy level.

Goodford's method is well suited for illustrating stable locations for the various functional groups which might be expected in ligand molecules. Therefore, the method may be helpful for introducing appropriate substituent groups to appropriate positions.

Ligand molecules usually have non-hydrophilic parts whose non-bonded interaction with protein atoms cannot be negligible. Goodford's method does not give one the total interaction energy including such non-specific interactions. Our method has an advantage to his method in giving the estimate of the total interaction energy between ligand and protein in real time.

On the other hand, Pattabiraman et al. have developed another docking method [27]. They approximated the parameters in the Lennard-Jones potential as the square root of the product of the atomic parameters of the interacting atom pair, and calculated two sets of grid point data which corresponded to the attracting and repulsive terms of the potential function.

Similar to our method, the method of Pattabiraman et al. could perform a real-time estimation of interaction energy during the interactive manipulation of a ligand molecule on a three-dimensional graphics, by using those three-dimensionally tabulated data. But for the calculation of the non-bonded interaction term, the algorithm used in our method is simpler and with less approximation from the exact atom pair type calculation than their method. Therefore, our method gives more accurate values within the limit of the grid point approximation.

The method of Pattabiraman et al. provided the molecular surface representation to show the environment inside the binding pocket during the docking process. But the molecular surface does not encode the geometry of the actual ligand atom positions as does our atom acceptable region representation, and it does not explicitly show the possible positions of hydrogen bonding ligand atoms.

The preparation of three-dimensional grid point data takes a rather long computational time. (In the dihydrofolate reductase example, it takes about 1 h on the VAX/780 and 3 min even on a HITAC S-810/20 supercomputer.) This may seem to be a shortcoming of our method. But once the grid point data have been prepared, one can use the data repeatedly while one is studying many ligand molecules on the same receptor molecule.

The estimated energy values displayed on the graphics device are updated within 1 s while we are applying rotation, translation or bond rotations to the ligand molecule. Most of the interval time is attributed to the data transfer time between the host computer and the graphics device, and computational time required for the energy estimation is negligible compared to the data transfer time. Such a data communication bottleneck will no longer exist on the IRIS workstation version of the program which is under development now.

One of the shortcomings of our grid method is discontinuity and errors in the estimated energy value due to the finite grid spacing. More accurate and continuous estimate of the energy value will be obtained if we use values interpolated from neighboring eight grid points.

Hydrogen bonding interaction plays an important role in molecular recognition processes in biological systems. Hydrogen bonds cannot be properly treated only with a combination of electrostatic interaction and non-bonded interaction, because they have different directionarity and range for different functional groups. Several potential functions for hydrogen bonds have been proposed which take into account not only the distance but also the angles around the hydrogen bonding groups [28, 29]. But such potential functions requires much more computation compared to non-bonded or electrostatic terms, and therefore cannot be used for macromolecular systems practically. Moreover, the problem becomes more difficult by the fact that the position of a hydrogen atom on a hydrogen bonding functional group cannot be exactly determined due to bond rotation of the functional groups. Almost all commonly used molecular mechanics programs do not use those strict hydrogen bond functions, but just use simple potential functions of distance between a pair of donor and acceptor atoms of a hydrogen bond.

The way of treating hydrogen bonding characteristics in our method has no precedent. We have adopted a rather empirical but effective method for estimation of hydrogen bonds instead of using the strict potential function. Taking into account distances and directions from lone-pairs or donor hydrogens in the macromolecule, expected spacial arrangement of opposing functional groups in the binding site cavity are illustrated by the use of the three-dimensional grid points. Our method may seem not quantitative as we do not use any potential function. But the criteria which we used to predict such favorable regions are reasonable from a statistical consideration.

In general, docking simulation methods have difficult problems about induced fit and treatment of water surrounding the protein. For the first problem, it has been shown that no significant conformational changes occur during the ligand docking in some protein series [30]. For the second problem, we may take an assumption as a primary approximation that only strongly bound water molecules found in the crystal structure analysis are significant and should be treated as a part of the protein structure.

CONCLUSION

Using our method, one can test binding of many ligand molecules to the same receptor in a short period, with the aid of real-time estimate of interaction energy and visualization of the environment of ligand binding site. Furthermore, our method will be of great use for modifying drug structures to optimize activity from a lead compound, and also for rationally designing new structures which can bind more strongly to the receptor protein.

Although our method can perform simple energy minimization of ligand geometry, molecular mechanics calculation may be necessary for more detailed study. Our method will be useful for finding initial geometry of ligand molecules before applying precise molecular mechanics calculations. By rationally determining the initial geometry of ligand molecule, the risk of entering into 'local minimum' during the molecular mechanics calculations will be reduced.

ACKNOWLEDGMENT

We thank Sankyo Co. Ltd. for the arrangement to use their computer graphics display and computer.

REFERENCES

1 Smith, R.N., Hansch, C., Kim, K.H., Omiya, B., Fukumura, G., Selassie, C.D., Jow, P.Y.C., Blaney, J.M. and Langridge, R., Arch.Biochem.Biophys., 215 (1982) 319–328.
2 Smith, G.M., Hangauer, D.G., Andose, J.D., Bush, B.L., Fluder, E.M., Gund, P. and McIntyre, E.F., Drug Inf. J., 18 (1984) 167–178.
3 Blaney, J.M., Weiner, P.K., Dearing, A., Kollman, P.A., Jorgensen, E.C., Oatley, S.J., Burridge, J.M. and Blake, C.C.F., J.Am.Chem.Soc., 104 (1982) 6424–6434.
4 Freudenreich, C., Samama, J.-P. and Biellmann, J.-F., J.Am.Chem.Soc., 106 (1984) 3344–3353.
5 Weiner, P.K., Langridge, R., Blaney, J.M., Schaefer, R. and Kollman, P.A., Proc.Natl.Acad.Sci. USA, 79 (1982) 3754–3758.
6 Blaney, J.M., Jorgensen, E.C., Connolly, M.L., Ferrin, T.E., Langridge, R., Oatley, S.J., Burridge, J.M. and Blake, C.C.F., J.Med.Chem., 25 (1982) 785–790.
7 Itai, A. and Tomioka, N., Abstracts of Papers, 6th Symposium on Medicinal Chemistry, Pharmaceutical Society of Japan, Tokyo, 1984, p. 105.
8 Tomioka, N., Itai, A. and Iitaka, Y., In Iitaka, Y. and Itai, A. (Eds.) Proceedings of the Symposium on Three-Dimensional Structures and Drug Design, Univ. Tokyo Press, 1987, pp. 186–194.
9 Goodford, P.J., J. Med.Chem., 28 (1985) 849–857.
10 Weiner, P.K. and Kollman, P.A., J.Comp.Chem., 2 (1981) 287–303.
11 Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S. and Karplus, M., J.Comp.Chem., 4 (1983) 187–217.
12 Weiner, S.J., Kollman, P.A., Case, D.A., Singh, U.C., Ghio, C., Alagona, G., Profeta, S., Jr. and Weiner, P., J.Am. Chem.Soc., 106 (1984) 765–784.
13 Warshel, A., J.Phys.Chem., 83 (1979) 1640–1652.
14 McCammon, J.A., Wolynes, P.G. and Karplus, M., Biochemistry, 18 (1979) 927–942.
15 Kollman, P.A., Weiner, P.K. and Dearing, A., Biopolymers, 20 (1981) 2583–2621.
16 Murray-Rust, P. and Glusker, J.P., J.Am.Chem.Soc., 106 (1984) 1018–1025.
17 Vedani, A. and Dunitz, J.D., J.Am.Chem.Soc., 107 (1985) 7653–7658.
18 Hansch, C., Li, R., Blaney, J.M. and Langridge, R., J.Med.Chem., 25 (1982) 777–784.
19 Kuyper, L.F., Roth, B., Baccanari, D.P., Ferone, R., Beddell, C.R., Champness, J.N., Stammers, D.K., Dann, J.G., Norrington, F.E.A., Baker, D.J. and Goodford, P.J., J.Med.Chem., 25 (1982) 1120–1122.

20 Ghose, A.K. and Crippen, G.M., J.Med.Chem., 27 (1984) 901–914.
21 Bolin, J.T., Filman, D.J., Matthews, D.A., Hamlin, R.C. and Kraut, J., J.Biol.Chem., 257 (1982) 13650–13662.
22 Matthews, D.A., Bolin, J.T., Burridge, J.M., Filman, D.J., Volz, K.W. and Kraut, J., J.Biol.Chem., 260 (1985) 392–399.
23 Koetzle, T.F. and Williams, G.J.B., J.Am.Chem.Soc., 98 (1976) 2074–2078.
24 Kosugi, N. and Kuroda, H., Chem.Phys.Lett., 74 (1980) 490–493.
25 Matthews, D.A., Bolin, J.T., Burridge, J.M., Filman, D.J. Volz, K.W., Kaufman, B.T., Beddell, C.R., Champness, J.N., Stammers, D.K. and Kraut, J., J. Biol.Chem., 260 (1985) 381–391.
26 Busetta, B., Tickle, I.J. and Blundell, T.L., J.Appl.Cryst., 16 (1983) 432–437.
27 Pattabiraman, N., Levitt, M., Ferrin, T.E. and Langridge, R., J.Comp.Chem., 6 (1985) 432–436.
28 Lippincott, E.R. and Schroeder, R., J.Chem.Phys., 23 (1955) 1099–1106.
29 Chidambaram, R., Balasubramanian, R. and Ramachandran, G.N., Biochim.Biophys. Acta, 221 (1970) 182–195.
30 Huber, R., Kukla, D., Bode, W., Schwager, P., Bartels, K., Deisenhofer, J. and Steigemann, W., J.Mol.Biol., 89 (1974) 73–101.