Automatic Identification and Manipulation of Receptor Sites in Proteins. 2. Electrostatic Complementarity Analysis for the Evaluation and Selection of Candidate Ligand Receptor Sites

Carlos A. Del Carpio, Yoshimasa Takahashi,* and Shin-ichi Sasaki

Department of Knowledge Based Information Engineering, Toyohashi University of Technology, Tempaku-cho, Toyohashi 441, Japan

Received October 27, 1992®

An analysis of the electrostatic characteristics of candidate receptor sites automatically identified from three-dimensional information for proteins is performed. On this basis, coefficients of complementarity of electrostatic potentials for the ligand–receptor system are derived. The analysis proves to be efficient for positioning a ligand molecule within a candidate receptor site and for deriving structural characteristics for molecules which bind in an active site. The information obtained in this way is relevant in assisting in computer-supported drug design processes.

1. INTRODUCTION

When a compound of high affinity for a receptor is available, in vitro experiments are carried out using an irradiated sample of the compound and in this way the receptor sites for that particular compound are located, identified, and even counted. The burden of the experiment is however enormous when there are several compounds of interest for which the information is necessary.

The sophistication of crystallographic techniques oriented to the analysis and elucidation of three-dimensional structures of proteins and the lower cost of computer power offer an alternative for tackling the problem. In our previous paper[1] we described a geometrical algorithm with which the analysis of the three-dimensional structure information of a protein allows the identification and location of all the clefts with the geometrical characteristics to host a specific ligand molecule in it.

The aim of the present work is to select regions within a protein where a given ligand molecule has a high probability of binding or interaction. In this study we introduce a novel method for modeling the underlying complementarity in such systems. The result is the selection of the pocket region most plausible to host a certain ligand from the set of pocket regions previously found from geometrical considerations. Accordingly, the analysis carried out here follows the identification of pocket regions in proteins that can host small molecules and the analysis of the geometrical and electrostatic complementarity of its components. Pocket identification is performed from geometrical considerations alone using crystal coordinates taken from the protein data bank (PDB). Affinities of the receptor site toward a certain ligand molecule are analyzed by means of the electrostatic potentials of the system (a ligand molecule is a compound known to interact with the receptor macromolecule from experimental or homological evidence).

In this paper we attempt to express the complementarity of electrostatic potentials of the ligand–receptor system and demonstrate how this complementarity plays an important role in the way the ligand nests in the receptor cavity of the protein.

The process is then used to analyze the characteristics of all the cavities identified from the geometrical constraints in order to select the most plausible to be host of the particular molecule.

To illustrate the adequacy of the method introduced here, we performed a computational experiment in which a methotrexate molecule is positioned in one of the receptors found in DHFR (dihydrofolate reductase) and is automatically identified. The position of the methotrexate molecule predicted by our method is in good agreement with that observed in the crystalline state of the system.

2. METHODOLOGY

2.1. Physicochemical Problem. The mechanisms of drug action are so complex that a drug design process may not be subject only to compounds that show strong binding ability to the receptor site. Nevertheless, a strong binding characteristic is the condition that must be fulfilled by a drug to generate high pharmacological activity.

The strength of this type of interaction can be adequately analyzed by comparing the change in free energy of the ligand before and after binding to the receptor. However, the complexity of the ligand–receptor system makes theoretical computations of this nature very limited in scope.

One way to analyze the interaction between ligand and receptor is by using the potential energy of intermolecular interactions. The energy of interaction expressed by the potential energy relation can be divided into two main terms:

$$E_{\text{tot.}} = E_{\text{vdw}} + E_{\text{el}} \tag{1}$$

the first term, E_{vdw}, is the van der Waals energy term. This term is relevant at small distances between two molecules and has the following form:

$$E_{\text{vdw}} = \sum_{i=1}^{n}\sum_{j=1}^{m}(A/d_{ij}^{12} - B/d_{ij}^{6}) \tag{2}$$

where d_{ij} is the distance between atoms i and j. A and B are constants which depend on the type of atoms involved in the interaction.

The second term, E_{el}, expresses the electrostatic interaction energy and has the following form:

$$E_{\text{el}} = \sum_{i=1}^{n}\sum_{j=1}^{m} q_i q_j / \xi d_{ij} \tag{3}$$

where q_i and q_j are the partial charges of atoms i and j and

* Abstract published in *Advance ACS Abstracts*, September 1, 1993.

0095-2338/93/1633-0769$04.00/0

ζ is the dielectric constant of the medium. Both terms reflect the geometrical as well as the electrophysical nature of the interaction between two chemical entities.

For the ligand–receptor system, the first term is important when the ligand is very close to the internal surface of the groove. The second one is a measure of the intermolecular electrostatic interaction.

It is then evident from the terms of the potential energy expression that the interacting components of ligand–receptor systems must have a complementary geometry as well as a complementary distribution of partial electrostatic charges.

Moreover, this suggests a methodology to select a cleft in a protein of known structural characteristics based on the position adopted by the ligand molecule in the cleft and the degree of electrostatic complementarity with its components.

2.2. Geometrical Complementarity. The geometrical aspect of ligand binding has been repeatedly treated in many researches. Crippen et al.[2] approached the problem from the topology of the system, finding cliques of the docking graph. The algorithm is designed to solve the combinatorial problem and presupposes knowledge of the interacting atoms.

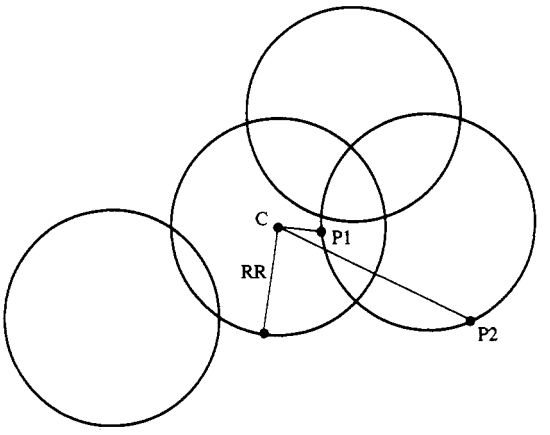
Although the algorithm succeeds in performing the task within a reasonable computational time, its complete disregard for the intrinsic physicochemical aspect of the interactions among the constituents of the system is incontestable.

Here we propose a technique that balances these aspects, i.e. the computer time and practicality of the solutions. It involves a minimization of the first term of the potential energy expression for many positions of the ligand molecule within a particular cleft (the candidate site identified by the system), and a subsequent analysis of the electrophysical complementarity of the components of the system when the ligand is placed in the optimal positions obtained by the minimization process.

The minimization operation performed here generates all the optimal positions that the ligand may adopt within a receptor cavity by minimizing the first term of the potential energy of the ligand–site, system. This is done by translations of coordinates along the three Cartesian axes in directions that lead to minimal van der Waals energy. The starting orientations are systematically derived by rotating the molecule by a determined increment in the rotational angles ϕ and θ. Computing time is proportional to n^2, n being the number of steps to describe a complete revolution of 360° in θ and 180° in ϕ.

A computer program based on the fast Fletcher and Reeves conjugate gradient minimization algorithm[3] yields all the optimal positions for the orientations in which the ligand molecule fits within the groove. These are the positions of minimal energy which the ligand molecule can adopt without colliding with the walls of the receptor for different orientations of the molecule (these positions are called "optimal positions" in what follows). The evaluation and final selection of the most plausible position among the set of optimal positions is done by analyzing the electrostatic complementarity of the system as described in the following section.

2.3. Electrostatic Complementarity. The position and orientation in which the ligand nests itself within the receptor cavity is, on the other hand, the result of the physicochemical characteristics of the interaction of both entities in addition to their corresponding geometrical complementarity. This is particularly the case when the geometry of the ligand is such that it can accommodate itself inside the cleft in different positions and orientations.



Exclusion of occluded atomic surface points.

$\overline{P(x,y,z)\,C} < RR \longrightarrow$ occluded point.(*)

$\overline{P(x,y,z)\,C} > RR \longrightarrow$ non occluded point.(*)

P1 ⟶ occluded; P2 ⟶ non occluded.

Figure 1. Computation of the nonoccluded surface points for the ligand molecule.

The present study is focused on one of the conformers of the ligand; the correlations derived in one operation are therefore associated with that specific conformer.

Analysis of the second term of the potential energy expression can be performed by computing the electrostatic potential of the ligand and that of the receptor site and comparing them. This comparison would yield the degree of complementarity of electrostatic potentials between both entities. A relatively high complementarity is expected to characterize the final position of the ligand molecule inside the receptor cleft.

A perfect complementarity of electrostatic potentials would be that in which both potentials have high numerical values but are of opposite sign at the point where they are compared.

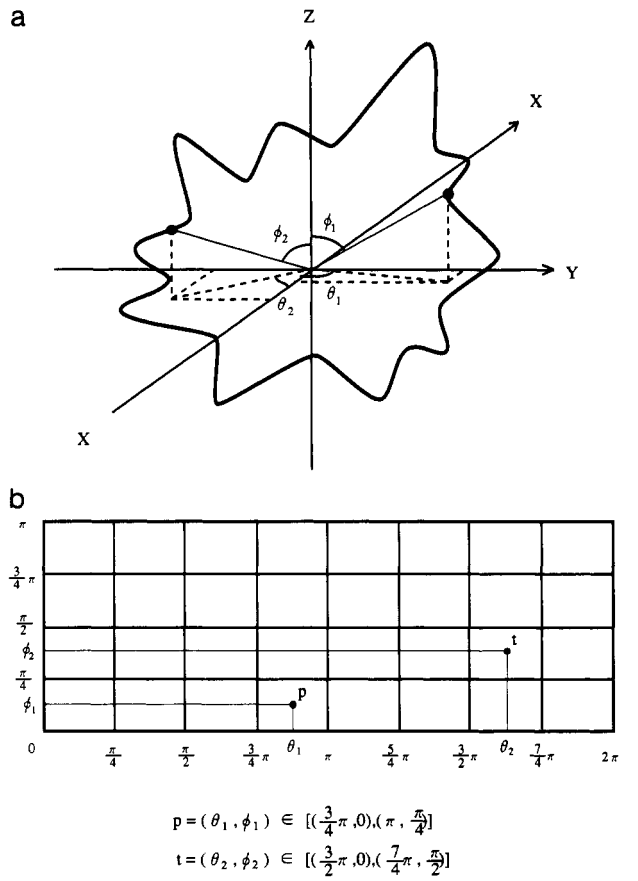
Several attempts to express electrostatic complementarity have been made. Nakamura et al.[4] defined a complementarity parameter for the electrostatic potentials due to the atoms of the protein and those of the ligand and reflected on the surface of the latter. This parameter however has the drawback that it is not representative at low values of the potential.

In this study electrostatic potentials of the protein are compared with the electrostatic potentials of the ligand in localized regions of the surface of the ligand molecule as described below.

The electrostatic potentials due to the atoms of the protein and those due to the atoms of the ligand are calculated and reflected on all the points of the surface of the atoms of the ligand molecule. These points are all nonoccluded points of the surface of the atoms (Figure 1). They are computed by circumscribing an icosahedral in a sphere representing the atom. The radius of the sphere is the van der Waals radius for each type of atom.

In order to perform the complementarity analysis, the coordinates of all nonoccluded or free surface points of atoms composing the ligand are transformed and projected as shown in Figure 2. Here, the Cartesian coordinates of the atoms constituting the cavity are transformed into polar coordinates (Figure 2a), taking as the origin of the polar coordinates system the center of gravity of the ligand molecule after geometrical optimization.

Electrostatic potentials for the protein site are calculated from partial electric charges for the atoms of amino acids



$$p = (\theta_1, \phi_1) \in [(\tfrac{3}{4}\pi, 0), (\pi, \tfrac{\pi}{4})]$$

$$t = (\theta_2, \phi_2) \in [(\tfrac{3}{2}\pi, 0), (\tfrac{7}{4}\pi, \tfrac{\pi}{2})]$$

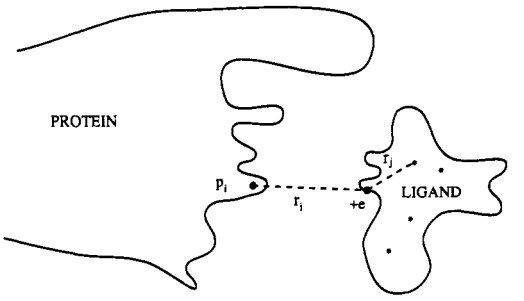
Figure 2. (a) Representation of the points forming the surface of nonoccluded points of the ligand molecule. (b) Regions of nonoccluded surface atoms.

derived from semiempirical MO calculations as described in ref 5. The electrostatic potential energy is calculated with the partial electric charges imposing a cutoff of 12 Å and a probe charge of $+e$ situated on the point of interest on the surface of the ligand (Figure 3).

Similarly, charges for the hypothetical ligand molecule are computed by the QCPE program AMPAC.[6] The electrostatic potentials for the ligand molecule are also calculated on the same surface as the electrostatic potentials for the protein atoms.

A global comparison of the electrostatic potentials calculated on all the points of the surface of the ligand distributed in the configuration introduced here yields a correlation coefficient of −0.401 for the system methotrexate–DHFR in the crystalline state extracted from the PDB (Figure 4). This value reveals a fair tendency to complementarity between both electrostatic potentials on the surface of the ligand molecule.

However, complementarities of electrostatic potentials are expected to be high in determined regions of the system while to a lesser extent in others due to the heterogenous distribution of partial charges on the surfaces of the atoms composing the ligand and the receptor groove. The examination of these type of localized complementarities would lead to a better assessment of the integrated complementarity of the system. Furthermore, it would also guide the process of identification and characterization of features regarding the composition and activity of plausible candidate ligands that could bind to a specific receptor. In order to take this aspect into consideration a regional correlation of the electrostatic potentials is more suitable. Accordingly, we introduce here a new procedure to evaluate the regional electrostatic complementarity for the system ligand receptor site. This



Electrostatic Potential for the Cavity

$$\psi_n = \sum_i^n \frac{q_i}{\epsilon r_i}$$

n = all the atoms of the protein within 12 Å around +e

r_i = distance from a protein atom i to the probe charge +e

$\epsilon = 2$

q_i = charge of atom i

Electrostatic Potential for the Ligand

$$\psi_m = \sum_j^m \frac{q_j}{\epsilon r_j}$$

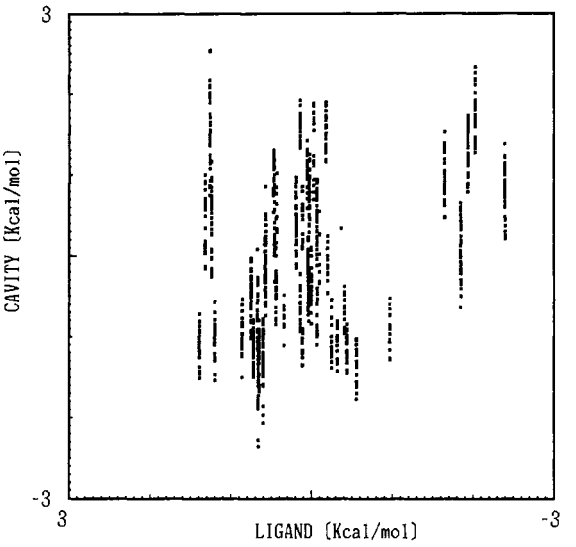
m = all the atoms of the ligand

r_j = distance from atom j of the ligand to the probe charge +e

$\epsilon = 2$

q_j = charge of atom j

Figure 3. Calculation of electrostatic potentials for the protein and ligand molecules and reflection on the surface of the ligand.



POSITION : T1= 0 ,P1= 0 T2=180 ,P2=360

Complementarity Analysis:

n= 1250 r= -0.401

Figure 4. Correlation of electrostatic potentials on the surface of the ligand.

procedure is based on the computation of correlation coefficients for regions composing the entire surface of the ligand molecule where the electrostatic potentials are calculated. To carry out this analysis, all nonoccluded points of the surface of the ligand are clustered in regions as shown in Figure 2b. In this study the whole set of nonoccluded surface points was divided in 32 regions or clusters, as illustrated in Figure 2b. Clusters of points belonging to a certain region can be expressed by a pair of values representing the angles of the leftmost point of the bottom and the rightmost point of the top of the region. For example, in Figure 2b, point $p = (\theta_1, \phi_1)$ belongs to region $[(3\pi/4,0), (\pi,\pi/4)]$.

Electrostatic complementarities are then computed for each localized group of points (region) of the surface of the ligand and for any of the optimal positions of the ligand of the groove obtained by the geometrical optimization process described above.

The complementarity analysis is performed by examining the regional complementarity between the electrostatic potentials for the atoms composing the protein receptor and the ligand molecule.

The map that results from displaying the regional complementarities on the surface at which the electrostatic potentials were computed is called the "complementarity space". This procedure is repeated for all the available optimized positions of the ligand within the cleft, and a complementarity space is obtained for each optimal position. The position which yields the best correlational coefficients or scores of complementarity is selected as the final optimal position that the ligand can adopt within the candidate receptor site, and finally the most plausible receptor site candidate.

The evaluation of a complementarity space on a region by region basis provides information on the strength of localized electrostatic interaction between the ligand and the groove that can characterize particular features related to the local geometry and structural composition of the system, whereas comparison of two complementarity spaces provides the means for selection of the most plausible to the receptor site. The former is performed by analysis of the local complementarity scores and the latter by analysis of local complementarity scores and comparison of the global correlation coefficient.

The analysis described above is exemplified by a computational experiment carried out to position the methotrexate ligand within one of the candidate receptor sites in the DHFR protein. The calculation of the probability of geometrical fitting of the molecule is performed by successively placing the ligand within each of the candidate receptors. The system assesses the possibility of geometrical fitting of the molecule within a cavity executing the geometrical analysis described above.

The optimization process yields the most probable orientations that the ligand can adopt within the receptor groove without colliding with the walls. Thereafter, the analysis of the electrostatic complementarity is carried out. Electrostatic potentials for the protein and the ligand are computed and reflected on all the nonoccluded points of the surface of the ligand. This set of points is divided into 32 regions, as explained before, complementarity coefficients are calculated for each region.

This calculation is repeated for all the geometrically optimized positions of the ligand molecule obtained by rotating it within the cavity.

For the present experiment an angle increment of 5° was selected. This equates to 2592 starting orientations for the geometrical optimization process. The results for the experiment are summarized in the next section.

3. RESULTS AND DISCUSSION

One of the aims of drug design experiments is to try to produce drugs with high affinity to the receptors, and this results in trials to construct putative candidate molecules with strong electrostatic complementarity with atoms bearing the highest charge and which are known to be forming a determined receptor site. This procedure leads to the wrong establishment of charges at site points for predicting complementary charges on the ligand that is being designed, expecting that the electrostatic potentials will somehow be complementary.

The wrongness of this assumption is evident from the fact that electrostatic potentials on the surface of the ligand cannot be derived only from some selected atoms conforming the cavity. This electrostatic potential is, as described by the potential energy equation, a distance-dependent summation of all the charges of the hull of atoms surrounding the ligand, and not merely those from a predetermined set.

Hence, the analysis introduced in the present work shows the extent to which treating the ligand–receptor system from the electrostatical point of view is appropriate in gauging the capability of a determined candidate pocket region to behave as a receptor site as well as the role that it can play in the determination of the binding position of a hypothetical ligand molecule within the receptor cleft.

The identification of the pocket regions in the protein provides the location of the groove where the protein can host a ligand molecule and hence the determination of the protein atoms which may directly interact with it. For the complex DHFR with NADPH the system assesses as candidate receptor site the cavity composed by the atoms listed in Table I. The analysis is accomplished by the geometrical optimization module. For the DHFR–methotrexate system, 43 optimal positions are obtained out of the 2952 starting orientations. These are positions where the van der Waals force term of the potential energy expression reaches a minimum, i.e. positions where the geometrical complementarity is achieved. Partial charges for the atoms of the amino acids are used to calculate the electrostatic potentials due to the protein, and the electrostatic potentials due to the ligand atoms are computed by using atomic charges obtained by means of a semiempirical molecular orbital method.

Here, the information of the three-dimensional structure of the conformer for methotrexate extracted from the crystal data was used for calculation of the partial charges. The potentials are calculated on the nonoccluded points of the surface of methotrexate as referred to before. Significant contributions to the value of the electrostatic potentials from the site are found for distances as far as 10–14 Å. Furthermore, a reasonable computation time is achieved when atoms located within this range of distances around the point of interest are considered. Therefore, a cutoff distance of 12 Å was used for the present analysis.

The regional or local complementarities are then computed for all the geometrically allowable positions and for all the regions on the surface of the ligand molecule. A visualization of the electrostatic potentials for the protein and the ligand is shown in Figure 5. Here the ligand molecule is shown inside the candidate cavity that is assessed by the system as host of the methotrexate molecule. The position of the ligand within the groove is that where the highest complementarity of electrostatic potentials is achieved by the calculations introduced in this study. The position of the receptor in the protein is also illustrated showing its location inside the DHFR protein whose representation is made by means of its secondary structure.

To systematize the results of the experiment, these are expressed by means of a change of orientation with respect to a referential position. For computation of this change in the position of the ligand within the cleft, the position of methotrexate in the crystal complex is taken as referential. The deviations of the atoms coordinates are expressed in terms of the root mean square (RMS) of the positions of the atoms of the rotated ligand and that of the referential crystal complex;

Table I. Constitution of the Methotrexate Host Cavity[a]

PN	Z	AT	NAA	AA	PN	Z	AT	NAA	AA	PN	Z	AT	NAA	AA
210	8	OD1	26	ASP	209	6	CG	26	ASP	211	8	OD2	26	ASP
50	6	CB	6	ALA	412	6	CD	50	PRO	249	6	CD2	30	PHE
246	6	CB	30	PHE	46	7	N	6	ALA	446	6	CD1	54	LEU
207	8	O	26	ASP	247	6	CG	30	PHE	447	6	CD2	54	LEU
34	6	C	5	TRP	47	6	CA	6	ALA	261	6	CZ	31	ARG
239	6	CE2	29	TYR	206	6	C	26	ASP	785	7	N	98	GLY
192	8	O	24	PRO	28	6	CB	26	ASP	229	7	NE2	28	HIS
394	8	OG	48	SER	222	6	C	28	HIS	33	6	CA	5	TRP
919	7	N	116	THR	445	6	CG	54	LEU	782	6	C	97	ALA
218	6	CD1	27	LEU	411	6	CG	50	PRO	251	6	CE2	30	PHE
825	6	CZ	103	PHE	31	6	CD2	4	LEU	221	6	CA	28	HIS
224	6	CB	28	HIS	405	6	CZ	49	PHE	780	7	N	97	ALA
454	6	CD	55	PRO	220	7	N	28	HIS	219	6	CD2	27	LEU
783	8	O	97	ALA	406	7	N	50	PRO	924	8	OG1	116	THR
243	6	CA	30	PHE	217	6	CG	27	LEU	142	7	N	19	LEU
228	6	CE1	28	HIS	259	6	CD	31	ARG	453	6	CG	55	PRO
193	6	CB	24	PRO	258	6	CG	31	ARG	35	8	O	5	TRP
920	6	CA	116	THR	367	6	CG2	45	THR	403	6	CE1	49	PHE
223	8	O	28	HIS	226	7	ND1	28	HIS	214	6	C	27	LEU
925	6	CG2	116	THR	149	6	CD2	19	LEU	36	6	CB	5	TRP
227	6	CD2	28	HIS	212	7	N	27	LEU	392	8	O	48	SER
473	7	NH1	57	ARG	262	7	NH1	31	ARG	775	8	O	96	ILE
28	6	CB	4	LEU	471	6	NE	57	ARG	470	6	CD	57	ARG

[a] PN, atom number in DHFR; Z, atomic number; AT, amino acid atom; NAA, number of amino acid; AA, amino acid.

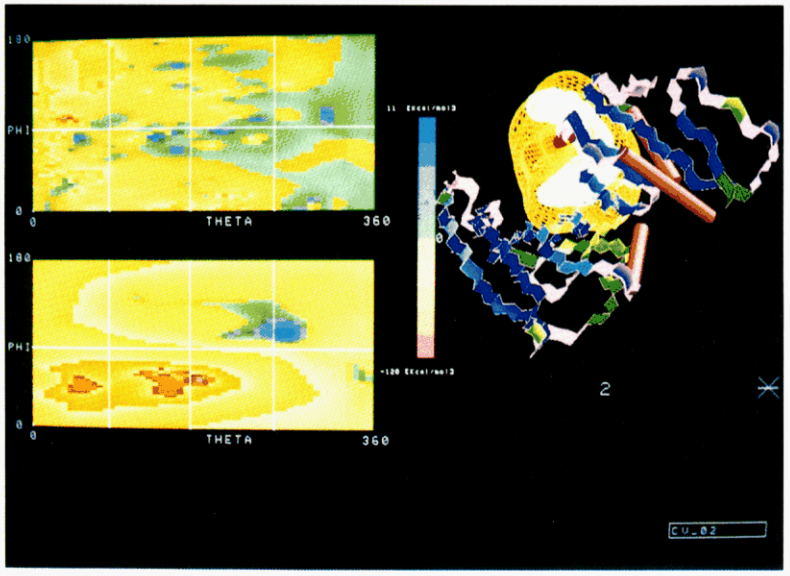


Figure 5. Visualization of the complementarity of electrostatic potentials for the components of the system receptor candidate of DHFR and the methotrexate ligand molecule.

its expression is

$$\mathrm{RMS} = 1/N \sum_{i=1}^{N} \{(x_i - x_{i0})^2 + (y_i - y_{i0})^2 + (z_i - z_{i0})^2\}^{1/2} \quad (4)$$

where x_i, y_i, and z_i, are the coordinates of atom i in the optimized position and x_{i0}, y_{i0}, and z_{i0} are the coordinates of atom i in the reference structure (i.e. in the crystalline complex). The rotation suffered by the ligand is also expressed (in this experiment) with regard to the referential position of the crystal and is expressed in degrees.

Table II illustrates the coefficients of the local correlations of the electrostatic potentials for the cavity and ligand for the 32 regions with the number of points for each region used in the correlation. It shows the complementarity coefficients for the best and the worst results out of the 43 different positions in which the ligand molecule can nest within the cavity together with coefficients calculated for the position that is found in the crystal structure. The first position (0), is that which the ligand is found to adopt in the crystalline form of the system. Positions (30) with rotation angles $\phi = 5°$ and $\theta = 35°$ are the optimized positions derived by the present analysis where the complementarity coefficients are higher or comparable to that of the crystal values. Position 43, with rotation angles $\phi = 175°$ and $\theta = 10°$, is the case where the complementarity coefficients are the lowest.

Figure 6 illustrates the degree of superposition of the 43 optimized positions resulting from the geometrical analysis with the position of the ligand in the crystal. The electrostatic complementarity analysis yields 14 structures having global complementarity scores similar or higher than the crystal structures.

Figure 7 illustrates the degree of superposition of the crystal position of methotrexate with the position of optimal com-

Table II. Optimal Positions for Methotrexate within the Receptor Cavity and Regional Coefficients of Electrostatic Complementarity[a]

NP	R	Np	R	Np	R	Np	R	Np	R	Np	R	Np	R	Np	R
Position 0: Rot. Ang., 0° and 0°; − −, 14; ++, 6; RMS, 0.000; rtot, −0.401															
0		153	−0.10	32	−0.31	8	0.00	8	−0.38	34	0.35	112	−0.10	12	0.03
0		0		1		0		0		105	−0.05	114	−0.20	4	0.05
129	−0.13	199	−0.25	0		0		19	−0.50	66	−0.27	0		0	
15	0.07	33	0.09	58	−0.08	2		39	−0.31	69	−0.46	37	0.41	1	
Position 31: Rot. Ang., 5° and 35°; − −, 14; ++, 8; RMS, 3.481; rtot, −0.431															
51	−0.04	132	−0.09	44	−0.24	1		7	0.57	63	0.14	93	−0.10	4	0.00
0		34	−0.19	0		0		12	−0.33	196	0.01	5	0.04	0	
11	0.00	200	−0.22	134	0.12	0		0		5		32	−0.44	0	
9	0.11	24	−0.42	38	−0.01	6	−0.54	15	−0.29	76	−0.35	55	−0.60	3	0.98
Position 43: Rot. Ang., 10° and 175°; − −, 5; ++, 17; RMS, 10.779; rtot, 0.338															
19	0.02	77	0.21	27	0.57	9	−0.68	66	−0.21	60	0.12	8	−0.44	0	
0		27	0.37	138	0.22	29	0.05	0		23	0.67	170	0.03	61	0.32
20	−0.26	165	0.16	0		0		0		0		0		5	
52	0.18	74	0.04	3	0.60	1	0.00	33	0.10	169	0.15	14	0.87	0	

[a] Abbreviations: Rot. Ang., rotation angles θ and ϕ, reference angles $\theta = 0$ and $\phi = 0$; − −, number of regions with negative complementarity coefficient; ++, number of regions with positive coefficients; RMS, root mean square of positions of atoms with respect to the referential position; rtot, global correlation coefficient for the optimal position. Np, number of points in each region used for the correlation analysis; *R*, correlation coefficient for the region.

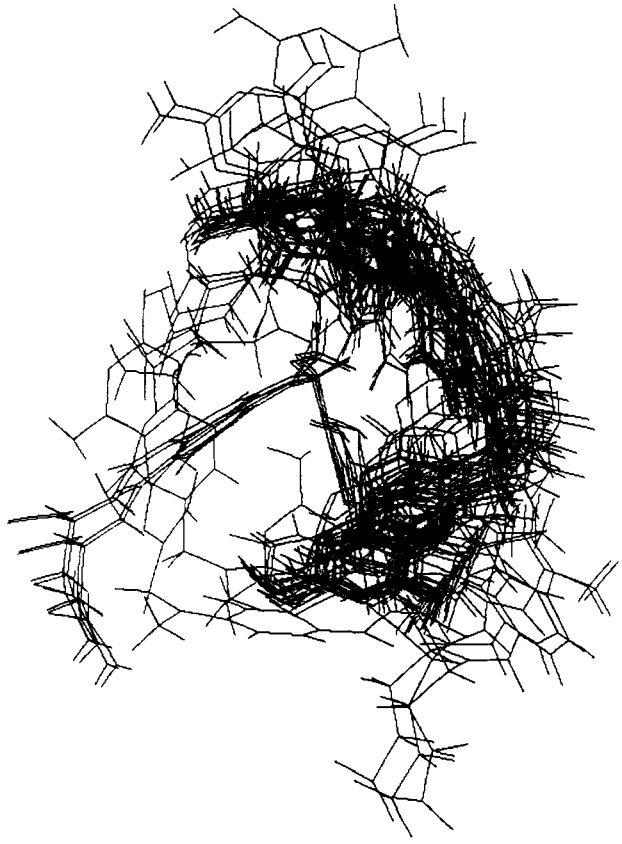


Figure 6. Superposition of the crystal conformer of methotrexate at the 43 optimized positions and its position in the crystal state.

plementarity (position 31 in Table II). Figure 8 illustrates the degree of superposition of the crystal structure with that of the position of lowest electrostatic complementarity (position 43 in Table II).

Most of the regional complementarity coefficients in positions close to the crystal position show a negative value, while most of the values are positive when the position of the molecule is completely changed.

Selection of the position with the highest degree of complementarity for the methotrexate–DHFR system is straightforward. The number of negative coefficients (i.e. regions where there is a marked tendency toward complementarity) and the global complementarity coefficient are the magnitudes in which the selection is based.

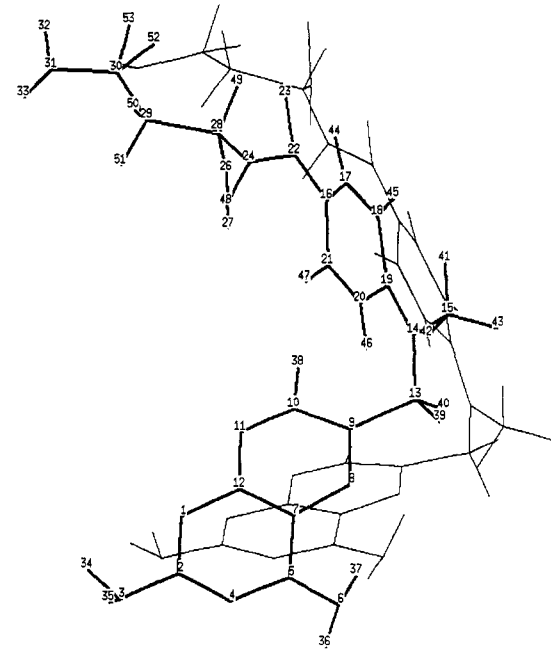


Figure 7. Superposition of the optimized position with high (no. 31 of Table II) electrostatic complementarity and the structure in the crystal (crystal structure in bold).

Positions that lie within 3.4 Å of RMS and rotated no more than 35° in ϕ and 10° in θ with respect to the referential position show a larger number of negative coefficients and are of higher absolute value than optimized positions that are completely different from the one adopted by the ligand in the crystal state.

When the ligand molecule is rotated more than 35° in ϕ and 10° in θ, and the RMS distance is higher than 3.5 Å, the number of positive local complementarity coefficients overpasses the number of negative coefficient regions, and correspondingly the global complementarity coefficient is positive or negative with very small absolute value (Table II).

For the methotrexate–DHFR system, the most plausible positions that the ligand can adopt are positions which are characterized by four regions where the local complementarity of electrostatic potential is high. These are the regions containing points of the nonoccluded surfaces of atoms listed in Table III.

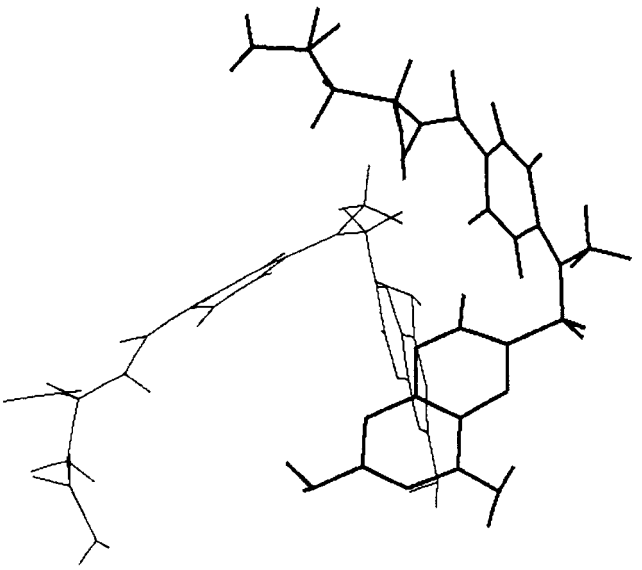


Figure 8. Superposition of the crystal conformer of methotrexate and the structure at optimized positions with lowest (no. 43 in Table II) electrostatic complementarity (crystal structure in bold).

Table III. Regions of Methotrexate Where High Electrostatic Complementary Tendency is Predicted[a]

region	composition	CC
[(5π/4,0),(3π/2,π/4)]	C16, C21, C22, C25, O27, O28	−0.35
[(3π/2,0),(7π/4,π/4)]	C17, C18, C22, O23	−0.60
[(3π/2,π/4),(7π/4,π/2)]	C31, O32, O33	−0.44
[(π,π/2),(5π/4,3π/4)]	N4, C5, N6, C7	−0.33

[a] composition: ligand atoms whose nonoccluded points conform to the region. CC: complementarity coefficient for the region (Table II, no. 31).

These are regions with high functionality in the methotrexate molecule to which its activity can be clearly attributed. The complementarity space mapped for the methotrexate in the DHFR site is illustrated in Figure 5 for the position with highest local complementarities that in this case correspond to a RMS of 3.48 Å, where only eight regions show a positive correlation coefficient, albeit their absolute value is small. The map is analogous to the map of the electrostatic potentials for the position in the crystallized complex. The outcome of placing the ligand molecule at the position where maximal regional complementarities of electrostatic potentials appear is in good agreement with the environments that the methotrexate molecule has in the crystal of the complex that was referenced from the PDB structural data file.

4. CONCLUSIONS

The analysis performed in this study shows the usefulness of including electrostatic interaction in the analysis of the complementarity of a ligand–receptor site to predict the hosting capability of the receptor, the binding capability of the ligand, and the optimal position that the ligand will adopt within the groove.

The candidate clefts have been automatically identified from the crystal structure information of proteins. Although direct comparison of the electrostatic potentials yields no obvious correlation on the surface of the ligand, regional correlations of electrostatic potentials perform better in portraying the intrinsically electrostatic complementarity of the receptor site and the ligand molecule.

Evaluation and comparison of the complementarity of electrostatic potentials for the different geometrically allowable positions of the ligand within the receptor groove lead to the selection of the most affinity enhancing position of the molecule within the groove.

Finally the selection of the receptor pocket region with the best affinity characteristics for the hypothetical ligand molecule is straightforward.

For the system methotrexate–DHFR, the system selects the region composed by atoms that are observed in the direct environment of the ligand in the crystal complex.

However, although the electrostatic potential energy for a rigid ligand is a significant term for the expression of the interaction patterns of ligand and receptor sites, a complete description of this sort of molecular interaction cannot neglect the flexibility of the components of the system. A further study that considers this factor is being performed in our laboratories, with the final aim of a global analysis of the interaction receptor ligand. Results of this study will be published in a future issue.

REFERENCES AND NOTES

(1) Del Carpio, C. A.; Takahashi, Y.; Sasaki, S. A New Approach to the Automatic Identification of Candidates for Ligand Receptor Sites in Proteins: (I) Search for Pocket Regions. *J. Mol. Graphics* **1993**, *11* (1), 23.
(2) Kuhl, F. S.; Crippen, G. M.; Friesen, D. K. A Combinatorial Algorithm for Calculating Ligand Binding. *J. Comput. Chem.* **1984**, *5* (1), 24.
(3) Fletcher, F.; Reeves, C. M. Function minimization by conjugate gradients. *Comput. J.* **1964**, *7* (2), 149.
(4) Nakamura, H.; Komatsu, K.; Nakagawa, S.; Umeyama, H. Visualization of electrostatic recognition by enzymes for their ligands and cofactors. *J. Mol. Graphics* **1985**, *3* (2), 11.
(5) Gruschus, J. M.; Kuki, A. Partial Charges by Multipole Constraint. Application to the Amino Acids. *J. Comput. Chem.* **1990**, *11* (8), 978.
(6) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. P. AM1: A New General Purpose Quantum Mechanical Molecular Model. *J. Am. Chem. Soc.* **1985**, *107*, 3902; *QCPE Bull.* **1986**, *6*, 506.