Protein Kinase C. Modeling of the Binding Site and Prediction of Binding Constants

Shaomeng Wang, George W. A. Milne,* Marc C. Nicklaus, Victor E. Marquez, Jeewoo Lee, and Peter M. Blumberg†

Laboratory of Medicinal Chemistry, DCT, and Laboratory of Cellular Carcinogenesis and Tumor Promotion, DCE, National Cancer Institute, NIH, Bethesda, Maryland 20892

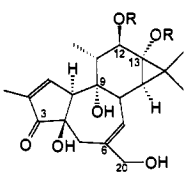
Received November 23, 1994®

A detailed examination of the mode of binding of phorbol esters to protein kinase C led to a model of the phorbol binding site in the enzyme. The efficacy with which various synthetic diacylglycerol analogs and ribonolactones are able to bind to this site was determined by means of semiempirical quantum mechanical calculations using PM3, and an estimate of the binding energy was made in each case. Sixteen synthetic analogs of 1,2-diacylglycerol and two natural products were studied, and their calculated energies of binding to this model were correlated with the measured K_i values. The binding energies calculated for this receptor model, together with solubility and entropy considerations, allow prediction through regressive fit of free energies of binding which correlate very well with the measured binding constants.

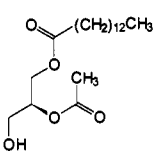
Background

Protein kinase C (PKC) plays a central role in cellular signal transduction, controlling diverse cellular processes including growth and differentiation.[1-6] The involvement of the PKC pathway in the promotion of oncogene action has focused attention on it as a promising target for chemotherapeutic intervention.[7] No X-ray crystallographic data have been reported for this enzyme, and studies of its mechanism of action have been impeded by this absence of data.

Numerous small molecules are known to bind to and activate PKC, and much contemporary knowledge of the regulatory active site of the enzyme where these compounds bind derives from studies of different PKC ligands. (Substances that bind to the protein are here referred to as ligands. The binding may be followed by activation of the enzyme which then interacts more effectively with its protein or peptide substrate.) Compounds in the 1,2-diacylglycerol family (DAGs) bind to PKC and activate it[1,2] while natural products such as esters of phorbol (1, R = H) also act as ultrapotent analogs of the DAGs.

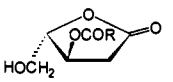


1



2

Binding of phorbol esters to PKC is an energetically favorable process; the inhibition constant (K_i) for unlabeled phorbol 12,13-dibutyrate (PDBU, 1, R = COC_3H_7) displacing the same radioactive compound from PKC is 0.000 78 μM.[8] Diacylglycerols such as 1-myristyl-2-acetylglycerol (GMA, 2) also compete with phorbol esters for the same site in PKC but have lower affinities—1.3 μM in the case of 2.[8] This may be attributable to the relative flexibility of 2, which suggests that, in contrast to the more rigid phorbol skeleton, this molecule may be forced to pay an entropy penalty when binding to the active site. Modifications, including cyclization, made to the DAG molecule did not vitiate its affinity for PKC; the lactone 3 was synthesized in this laboratory and shown[9]



R = $CH_3(CH_2)_{12}$-

3

to bind to PKC with a binding constant of about 2.5 μM. Binding to the regulatory site in PKC is known[10-17] to involve the three-point pharmacophore shown in Figure 1. This is made up of the ketone oxygen at C_3 and the hydroxyl groups at C_9 and C_{20} and augmented by a lipophilic residue such as a fatty acid side chain at C_{12}. The DAG has all the necessary binding points but, perhaps because it is so flexible, it binds less successfully to the active site. This information represented a starting-point for this paper which describes work designed to characterize the active site of PKC in greater detail.

Several approaches have been developed to model possible ligand binding sites. These include the hypothetical active-site lattice,[18] Crippen's distance geometry[19] and Voroni binding sites[20,21] methods, comparative molecular field analysis (CoMFA),[22,23] Goodford's GRID method,[24] and the RECEPS method.[25] In this paper we report on the use of pharmacophore geometry together with considerations of entropy and solubility to arrive at theoretical estimates of enzyme affinities.

Methods

1. **Molecular Modeling.** All molecular mechanics manipulations were carried out on a Silicon Graphics[26] IRIS Indigo workstation using Quanta.[27] Energy calculations and minimizations were carried out with the semiempirical quantum-mechanical method PM3[28,29] in the MOPAC 6.0 package[30] on a host mainframe Convex C240 supercomputer under Convex OS 10.2 (based on BSD Unix 4.2). In this arrangement, Quanta, which has good graphics management software, was used as a front end to the PM3 program, which is essentially noninteractive. An accurate method of assessing hydrogen bonds is clearly required for this work, and it was therefore necessary to select a computational method that is capable of accounting for hydrogen bonds. High-level *ab initio* methods in, for example, Gaussian or GAMESS are considered to deal well with hydrogen

† Laboratory of Cellular Carcinogenesis and Tumor Promotion.
* Abstract published in *Advance ACS Abstracts*, March 1, 1994.

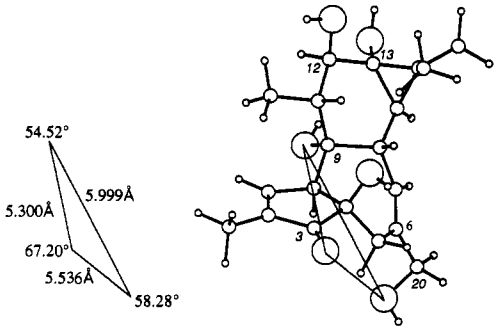


Figure 1. Structure of phorbol derived by addition of hydrogens to the X-ray-determined structure (PORBET10). The three-point pharmacophore is outlined and its dimensions are given.

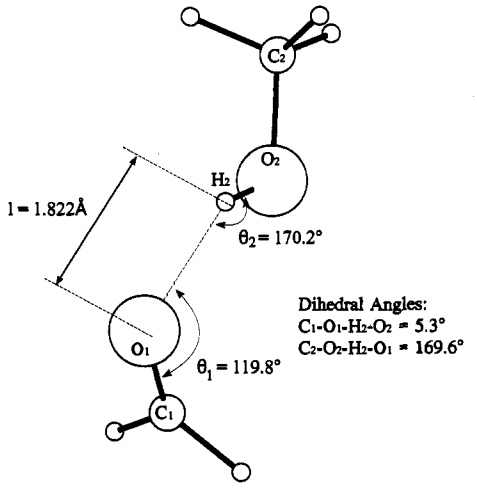


Figure 2. Global minimum energy structure of the methanol–formaldehyde system.

bonds, but for large molecules, they require more computer power than can conveniently be provided. At the other end of the spectrum, molecular mechanics methods fail to predict hydrogen-bond characteristics correctly. The semiempirical methods such as PM3, MINDO/3, MNDO, and AM1 represent a compromise between these two extremes. PM3 in particular is known[31] to calculate hydrogen bonds quite well, predicting bond lengths with errors of less than 0.2 Å and hydrogen bond energies with errors of about 1 kcal/mol. Accordingly, this program was used here. All PM3 minimization was carried out using the keyword "PRECISE".

Calculation using PM3 on a simple test system, the methanol–formaldehyde bimolecular complex, provided support for the decision to use this program in the pharmacophore measurements. Minimized structures of both formaldehyde and methanol were placed in proximity to one another, and the system was fully minimized. The hydrogen-bond structure that was obtained reproducibly from this experiment is shown in Figure 2. The hydrogen bond between the methanol proton and the formaldehyde oxygen has an optimum length of 1.822 Å and the system is nearly linear, the O_1–H_2–O_2 angle in the energy-minimized system being 170.2°. This is consistent with many other measurements[32] of hydrogen-bond dimensions and provides some justification for the choice of PM3 in the pharmacophore measurements.

2. Bioassay of PKC Binding. The binding affinity of different ligands for the phorbol ester site of PKC was measured by competitive displacement of [20-^{3}H]PDBU (1, R = COC_3H_7) from the enzyme by the compound being assayed. The method used has been described previously.[9]

Results

A. Development of the Binding-Site Model. 1. PDBU as a Template for the Regulatory Site of PKC. a. The Template Molecule. Structural data for PKC, in particular its binding site, are not available, but phorbol 12,13-dibutyrate (PDBU, 1, R = COC_3H_7) is known to bind very effectively to PKC, and since the structure of phorbol itself has been established by X-ray crystallography,[33] PDBU was chosen as the template molecule to model the binding site of PKC. The phorbol nucleus is essentially inflexible with the exception of the hydroxymethyl side chain at C_6, and it was assumed that the conformation derived from the crystallographic study is, with the possible exception of this side chain, the conformation in the enzyme–ligand complex.

b. Basic Phorbol Ester Atomic Coordinates. An X-ray crystallographic determination[33] of the structure of phorbol provided a starting set of atomic coordinates for the molecule, and these coordinates were used in the modeling programs. Since the X-ray study did not provide the atomic coordinates for the hydrogen atoms, hydrogens were added to the heavy atoms. The complete structure was then energy minimized with PM3 to establish the optimum positions of the hydrogens. In this way, the structure shown in Figure 1 was arrived at and used as the starting point for this work.

The minimized structure obtained in this way was compared to the original X-ray data. Superimposing all the non-hydrogen atoms of both structures gave an RMS value of 0.136 Å; the fit for the 3-point pharmacophore (C_3=O, C_9O, and C_{20}O) was 0.061 Å, and for the two-atom match involving C_3=O and C_9O, the RMS value was 0.017 Å. These data showed that the PM3-minimized structure of phorbol agreed well with the structure determined by X-ray crystallography, the positions and the orientations of the pharmacophore atoms being essentially unchanged by the minimization. The conformation of the C_6–C_{20}–OH in the bound PDBU is not known with certainty because the structure established by crystallography is, with respect to this hydroxymethyl side chain, only one of the several local energy minima.

c. The Pharmacophore of Phorbol Ester in PKC Binding. Previous studies[10-17] have shown that the binding of phorbol esters to PKC essentially involves three of the phorbol atoms, namely the oxygen of the carbonyl group at C_3 and the oxygen of the C_9 hydroxyl, both of which behave as hydrogen-bond acceptors, and the hydroxyl group at C_{20}, which is a hydrogen-bond donor. It follows that in the receptor there must be at least two hydrogen-bond donors complementary to the oxygens at C_3 and C_9 and a hydrogen-bond acceptor, corresponding to the C_{20} hydroxyl group. The spatial positions of these three atoms in the receptor must be such that strong hydrogen bonds are formed to phorbol, whose esters bind very effectively to the receptor. (The 12,13-esters of phorbol all bind to PKC, but phorbol itself fails to do so, presumably because of its poor lipid solubility. The issue of solubility is discussed later in this paper.) Thus from the structure of phorbol and the geometric characteristics of the optimum hydrogen bond, several possible alternative structures for the receptor can be deduced.

2. Modeling of the Active Site. a. Structures of Three Alternative Receptor Models. The three-point pharmacophore of phorbol based upon X-ray crystallographic data as described earlier was used as the template for modeling the PKC binding site. For the model, the atoms of the receptor which are involved in hydrogen bonds with the crucial phorbol atoms were modeled by very simple "probe" molecules. Methanol was used as the

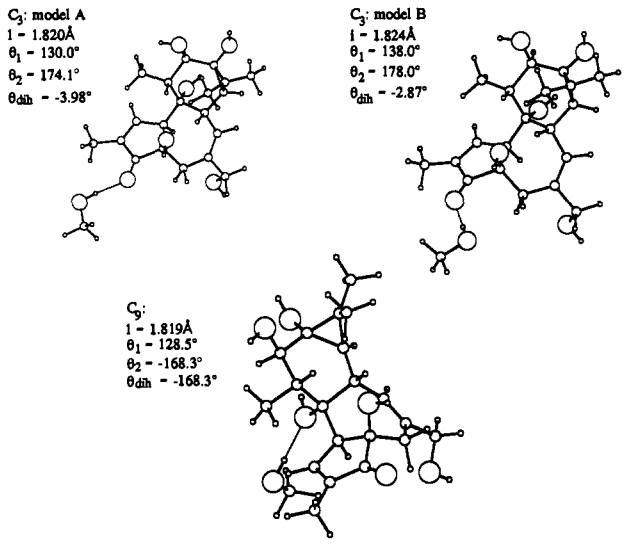


Figure 3. Optimum approach of hydrogen-bond donor (methanol) to oxygen at C_3 (models A and B) and at C_9 of phorbol (see Figure 2 for definitions of bond lengths and angles).

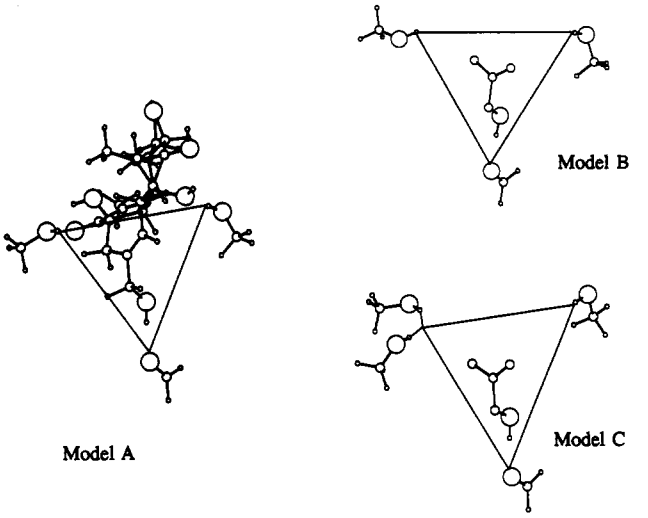


Figure 4. Models A–C of the three-point PKC receptor. Model A contains a molecule or phorbol imbedded in it for illustrative purposes. Models B and C show the hydroxymethyl side chain of phorbol.

hydrogen-bond donor and formaldehyde as a hydrogen-bond acceptor at the binding site.

The oxygen of the C_3 carbonyl group of phorbol has two lone pairs of electrons, either or both of which could become involved in a hydrogen bond to a second molecule. These are at approximately 120° to one another, and a hydrogen-bond donor such as MeOH can, as depicted in Figure 3, approach from two directions (models A and B). Therefore, for this binding site, the PKC receptor can, in principle, form one hydrogen bond from either direction, or form two hydrogen bonds by interacting with the two lone pair electrons simultaneously. These two hydrogen bonds were energy minimized with PM3, and the results, labeled "model A" and "model B", are shown in Figure 3. The PM3 calculations show that the hydrogen bond formed in the A direction has an energy of about 2.7 kcal/mol, and in the B direction, it is 3.5 kcal/mol. When both hydrogen bonds are formed simultaneously, the stabilization energy is 6.1 kcal/mol. It is interesting to notice that although there is only one hydrogen bond formed in A or B, B is more stable than A by 0.80 kcal/mol. This extra energy is due primarily to the hydroxyl group at C_4 which forms an additional weak hydrogen bond with the methanol. It is of interest in this connection to note that absence of the 4-hydroxyl group from the phorbol structure has been associated[34,35] with decreased biological potency, suggesting that this hydroxyl group may augment the three-point pharmacophore, although its presence is not obligatory. The optimal approach of a methanol to the oxygen at C_9 was determined in the same way, and the resulting model is also shown in Figure 3.

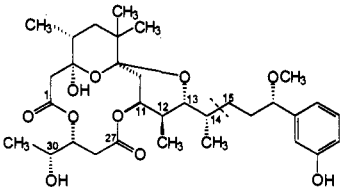
The third binding point in the phorbol molecule is the primary hydroxyl group at C_{20}. This functions as a hydrogen-bond donor, and so to model the interaction at this point, a typical hydrogen-bond acceptor, formaldehyde, was used in juxtaposition to the C_{20} hydroxyl group. In order to determine its precise position, the location and orientation of the hydroxyl group at C_{20} of phorbol must be determined. Examination of the torsion angles of the C_6–C_{20} bond in phorbol revealed its optimal value to be 73.6°. The bond between C_{20} and the O atom, however, rotates quite freely; a conformational search of this bond revealed two local minima differing by only 0.4 kcal/mol and separated by a barrier of 1.3 kcal/mol. In view of these results, it was felt to be prudent to allow free rotation of the C_{20}–O bond in the modeling study.

The location of the oxygen attached to C_{20} was thus fixed but its *orientation*, i.e., the torsion angle of the C_{20}–O bond, was allowed to vary. The optimum position of the hydrogen bond acceptor—the formaldehyde—in the model of the receptor was determined by energy minimization of a phorbol–formaldehyde complex which was found to be have dimensions almost identical to those obtained from the formaldehyde–methanol system shown in Figure 2. In subsequent enzyme–ligand fits, this hydroxyl group is allowed to rotate, and as it does so, the formaldehyde molecule tracks it so as to maintain its most effective interaction, i.e., lowest-energy orientation with respect to the phorbol hydroxyl.

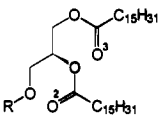
Three possible receptor models A–C were developed by combining the three binding sites on the receptor that correspond to the three oxygens of the phorbol molecule. They are shown in Figure 4. Models A and B have three probe molecules, two methanols and one formaldehyde, while in model C, there are four probe molecules, three methanols and one formaldehyde. Two of the three methanols in model C approach the oxygen of the C_3 simultaneously, forming two hydrogen bonds, a third methanol approaches the oxygen at C_9 to form another hydrogen bond, and the formaldehyde approaches the C_{20} hydroxyl to form the fourth hydrogen bond. These three models represent the most likely possibilities for the structure of the active site, and the next step is to determine which of these can interact most effectively with compounds which are known to bind to PKC and which can therefore be of predictive value.

b. Evaluation of the Alternative Receptor Models. Evaluation of models A–C was accomplished by comparing the measured binding affinities of a number of known PKC activators or agonists with the calculated binding energies to see which one of the models reproduces the experimental data best.

(i) Compounds Studied. The PKC ligands chosen included debromoaplysiatoxin (DAT, **4**), 1,2-dipalmitoyl-*sn*-glycerol (**5a**), and six rigid synthetic bicyclic lactones (**6**–**11**) which were previously designed and synthesized[39] in this laboratory.



4



5a R=H
5b R=CO1CH3

Debromoaplysiatoxin. Debromoaplysiatoxin (4), a derivative of the natural product aplysiatoxin A, has been shown[16] to bind very effectively to the regulatory site of PKC, with a K_i value of 0.02 μM. It is known that debromoaplysiatoxin competes with phorbol esters, which suggests they bind to the same receptor site. The structure of debromoaplysiatoxin was obtained from X-ray crystallographic measurements[36] and is shown in Figure 5. As can be seen, the phenolic side chain of debromoaplysiatoxin is quite flexible and folds back toward the hydrophilic surface of the molecule. This phenolic chain is, however, irrelevant to the binding energy calculation, and so for the modeling experiments, it was replaced by a hydrogen at C_{14}. Hydrogen atoms, their precise locations determined by energy minimization, were added to the X-ray-derived structure. The three-point pharmacophore of the debromoaplysiatoxin molecule is comprised of the oxygens at C_1 and C_{27} and the hydroxyl group at C_{30}, which correspond to the oxygens at C_3, C_9, and C_{20} of phorbol, respectively.[16] The PM3-minimized structure of the debromoaplysiatoxin ring system was compared to that in the X-ray determined structure. For all the heavy atoms, the RMS difference between these two structures was 0.307 Å. For the pharmacophore atoms, the RMS difference was 0.184 Å and the RMS between the two carbonyl oxygens of each structure was 0.177 Å. Thus although the macrocyclic ring in debromoaplysiatoxin is not as rigid as the tetracyclic ring system of phorbol, the pharmacophore atoms in the debromoaplysiatoxin structure were not significantly displaced by the minimization process.

1,2-Dipalmitoyl-*sn*-glycerol. Cellular PKC is activated when diacylglycerols bind to its active site, and it is therefore desirable to evaluate the receptor models using a diacylglycerol structure. The 1,2-diacylglycerols are generally very flexible, having up to seven rotatable bonds between the three crucial oxygens, the hydroxyl oxygen, and the two acyl carbonyl oxygen atoms. They can adopt many low-energy conformations, and the published crystal structures of 1,2-diacylglycerols cannot be used, because when bound to the enzyme, these compounds are likely to have a different conformation than in the crystal state. In 1,2-diacyl-3-acetylglycerols on the other hand, intramolecular hydrogen bonding is largely absent, and as a result, these compounds adopt a conformation that is quite different from that of the crystalline 1,2-diacylglycerols and which is not changed when the compound is in solution. NMR studies have shown[37] that the conformation of the fully esterified glycerols is very close to that of the 1,2-diacylglycerols in the lipid bilayer. The conformation of 1,2-dipalmitoyl-3-acetyl-*sn*-glycerol (**5b**) in the crystalline form has been determined[38] and is shown in Figure 6. It is very similar to the conformation adopted by diacylglycerols in the bilayer and is therefore considered[38] to be the conformation appropriate to diacylglycerol binding to PKC and should have a strong interaction with at least one of the three models. The crystal structure of **5b** was imported into Quanta. The acetyl group attached to the

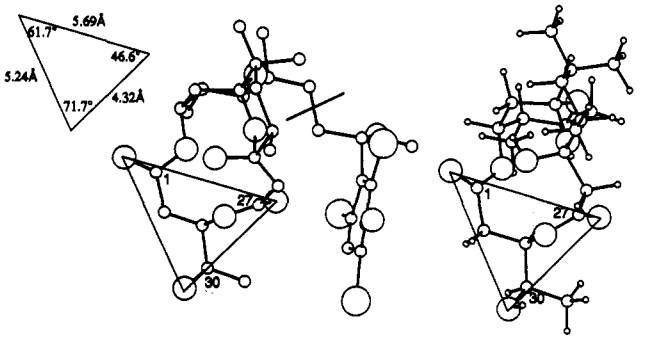
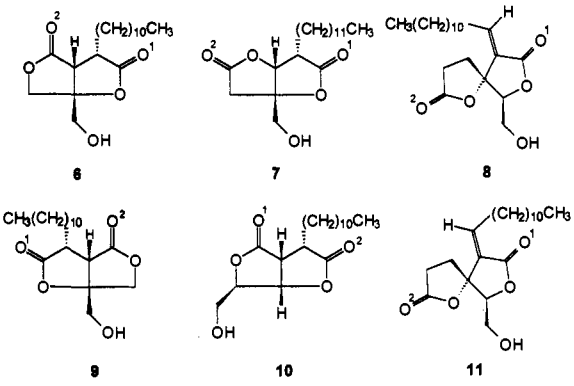


Figure 5. Structure of debromoaplysiatoxin (FUBXIV) from X-ray crystallography (left) and after removal of the side chain, addition of hydrogens, and minimization (right). The three-point pharmacophore and its dimensions are given.

oxygen at C_3 in the triglyceride was replaced by a proton, the long-chain acyl residues were changed to acetyls, hydrogens were added to other heavy atoms, and the structure was optimized using PM3, giving the second structure shown in Figure 6, which was used in subsequent experiments.

Rigid Synthetic Lactone Analogs. A number of bicyclic or bis-γ-butyrolactones (6–11) containing the 1,2-diacylglycerol moiety have been synthesized in this laboratory[39] and found to compete with phorbol esters for the binding site in PKC, which suggests that they bind on the same binding site of PKC as does phorbol 12,13-dibutyrate. The measured phorbol binding inhibition constants for debromoaplysiatoxin, the diacylglycerol (**5a**), and the six lactones are shown in Table 1. The ring systems of these lactones, minimized using PM3, are all relatively rigid, and it is assumed that, as a result, the relative positions and orientations of the two crucial carbonyl oxygens will not be significantly modified in the binding process.



6 7 8 9 10 11

(ii) Mapping of the Pharmacophore in Lactones 6–11. Pharmacophore mapping—the detection and characterization of a pharmacophore in a putative ligand—is used widely in drug design and was applied to the PKC binding problem. As a starting point, the pharmacophore is the three-point model derived from phorbol (Figure 1), and mapping to this model of the corresponding three-atom groups in the various test structures 6–11 was carried out by geometry and energy analyses.

Geometry Fit. Table 1 lists the RMS values derived from both two-point and the various possible three-point fits between phorbol and debromoaplysiatoxin (4), **5a**, and the six synthetic lactones (6–11). Table 1 reveals a number of discrepancies between the root-mean-square (RMS) values and the measured inhibition constants. For

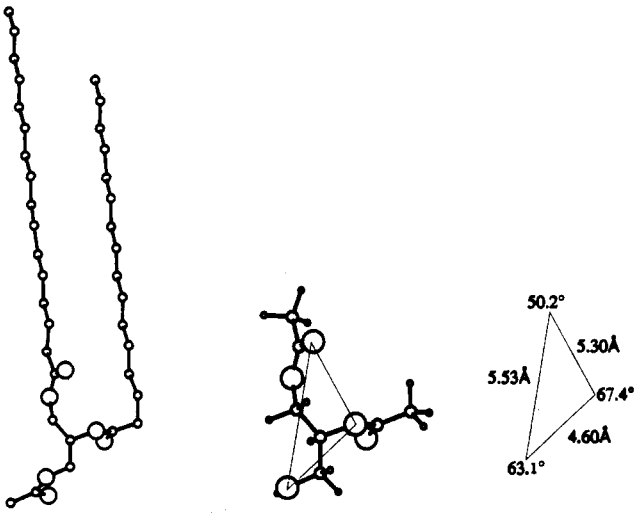


Figure 6. Structure of a triacylglycerol (left) and the modeled diacetylglycerol (right). The pharmacophore and its dimensions are given.

Table 1. Geometric Fitting of Phorbol Inhibitors to the Phorbol Three-Point Pharmacophore

compd no.	RMS fit (Å) two-point	three-point	three-point[a]	measured K_i (μM)
4	0.214	0.366	0.375	0.02
5a	0.351	0.492	0.362	[1.3][b]
6	0.242	0.313	0.525	7.7
7	0.098	0.261	0.477	7.6
8	0.146	0.191	0.183	13.6
9	0.246	0.246	0.520	22.0
10	0.285	0.329	0.399	27.8
11	0.048	0.266	0.139	75.3

[a] There are two possible orientations of the 3-point fit (see text). [b] The K_i value for GMA (2) was used here.

example, it can be seen that although debromoaplysiatoxin, with a K_i of 0.02 μM, binds with by far the strongest affinity, it does not give the lowest RMS value as might be expected. Some of the lactones give better fits (lower RMS values) to the phorbol pharmacophore. For example, lactone 11 has an RMS value of 0.048 Å in the two-point fit and 0.266 and 0.139 Å in the three-point fits, but showed low binding affinity with a K_i value of 75.3 μM in the PKC binding assay. In general, the lactones all showed only low-to-moderate binding affinities, considerably less than the affinity of debromoaplysiatoxin. This is inconsistent with their good fit to phorbol—a good fit which supported their original design.[39] The low affinity of the lactones for the active site cannot be attributed to an entropy effect since the lactones are no less rigid than debromoaplysiatoxin and phorbol. If the low binding affinity of the lactones is not attributable to entropy, it must result from enthalpy, which in turn suggests relatively poor interactions, not revealed by the simple geometry fit, between these compounds and the phorbol pharmacophore.

These results suggest that the simple pharmacophore model is unable to distinguish reliably between inhibitors with high binding affinity and those with low affinity. The reason for this appears to lie in the fact that superpositioning takes account only of the position of an atom, not of its orientation. Position alone is a necessary but insufficient criterion for binding. An oxygen in the correct position, but with the wrong orientation, particularly of its lone-pair electrons, will be unable to form a strong hydrogen bond with a hydrogen-bond donor—in this case the external methanol. This deficiency can be expected to cause a serious problem with attempts to use superpositioning on this simple pharmacophore model to design new inhibitors with high binding affinity, and explicit calculation of binding energies is necessary to solve this problem.

Two-Point Binding Energy Calculations. The binding energy will be a major determinant of the binding constant, although the entropy, which will contribute to the free energy of binding, is a secondary factor and is considered below. As an intermediate step in building the full, three-point receptor model, a preliminary two-point binding model was developed.

The two-point binding energies for molecules 4, 5a, and 6–11 in models A–C were determined directly from the models in an attempt to correlate binding energies with the measured binding affinities. As is discussed below, the ligand can bind to each model in two orientations, and these model/orientation combinations will be referred to here as A_α and A_β, B_α and B_β, and so on, with the orientations themselves designated as α and β. All of the compounds except the diglyceride 5a are relatively rigid, and the relative positions and orientations of the two carbonyl oxygens should not change during the binding process. With rigid structures of this sort there should be few entropy-related problems, and it is reasonable to assume that lack of high binding affinities for these lactones was caused by the weak interactions between the two carbonyl groups and the hydrogen bond donors in the receptor. To substantiate this point, binding-energy calculations were performed between these compounds and the two binding sites (corresponding to C_3 and C_9 of phorbol) which function as hydrogen-bond donors in models A–C.

The binding energies between debromoaplysiatoxin (4) and these three models, A–C (see Figure 4), were calculated by inserting the debromoaplysiatoxin into the model receptor, ensuring that the C_1 and C_{27} oxygens of debromoaplysiatoxin were both approximately 1.8 Å from the protons of the hydroxyl groups of the model's methanol molecules and that debromoaplysiatoxin's C_{30} hydroxyl group was within 3 Å of the oxygen of the formaldehyde. The formaldehyde was used only to ascertain that the orientation of the substrate was consistent with a three-point fit. To obtain the (two-point) binding energy between the debromoaplysiatoxin and the two methanol binding sites in the model, the formaldehyde was removed from this system, and the resulting complex, which contains debromoaplysiatoxin together with two (models A and B) or three (model C) methanol probe molecules, was submitted to PM3 for minimization. In the minimization process, the positions of the probe molecule atoms were kept fixed and the debromoaplysiatoxin was allowed to move in space to find the position in which it can form the strongest two-point interactions with the probe molecules. The conformation of the debromoaplysiatoxin was kept fixed. After minimization, the energy (enthalpy) of the complex formation at 298 K ($E_{complex}$) was obtained.[40] The energy of the ligand in the complex (E_{ligand}) was obtained by removing all the methanols from the complex and recalculating the energy of the system without geometry minimization. [The unqualified word "energy" is used here to signify enthalpy (ΔH). When free energy is discussed later, it is identified as "free energy" (ΔG).] Similarly, after removal of debromoaplysiatoxin from the complex and recalculation of the energy, without geometry

Table 2. PM3-Calculated Two-Point Binding Energies (kcal/mol) of Compounds with Different Models of the PKC Receptor[a]

compd no.	model A_α	model A_β	model B_α	model B_β	model C_α	model C_β
1 (R = H)	-5.72	bump	-6.50	bump	-9.70	bump
4	-6.50	-2.66	-6.44	-4.42	-9.33	bump
5a	-5.66	-4.65	-5.40	-3.99	-7.66	-6.06
6	-3.89	-3.20	-3.99	-4.24	-6.39	-4.47
7	-4.72	-4.44	-4.10	-3.89	-6.39	-4.79
8	-3.89	-3.00	-4.88	bump	-6.61	-5.35
9	-3.45	-2.33	-4.56	-3.39	-6.29	-4.05
10	-4.70	-3.89	-4.01	-3.89	-5.75	-3.98
11	-3.89	-3.00	-4.88	bump	-5.98	-3.93

[a] "Bump" means collisions occurred between the model and the substrate molecule in this binding mode, preventing formation of effective hydrogen bonds.

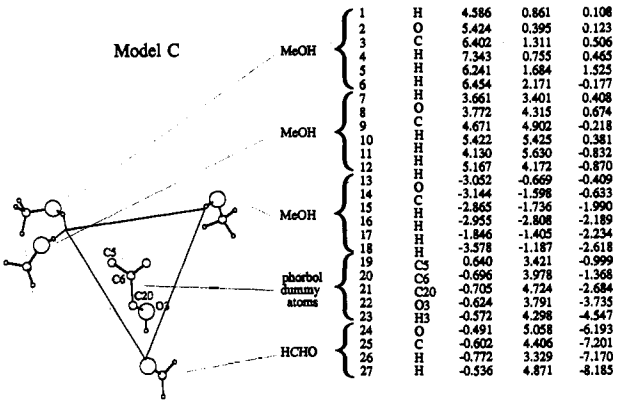
minimization, the energy of formation at 298 K of the model (E_{model}) is obtained. The binding energy ($E_{binding}$) was then calculated using the following equation:

$$E_{binding} = E_{complex} - (E_{model} + E_{ligand}) \quad (1)$$

The results of these calculations are shown in Table 2, in which it can be seen that, for both models, the binding energies of debromoaplysiatoxin (4) in the first orientation (models A_α [-6.50 kcal/mol] and B_α [-6.44 ckal/mol]) are substantially larger than in the other orientation (A_β [-2.66 kcal/mol], B_β [-4.42 kcal/mol]). It follows that α is the better orientation of debromoaplysiatoxin, in which the C_1 oxygen of debromoaplysiatoxin corresponds to the C_3 oxygen or phorbol, while the oxygen at C_{27} in debromoaplysiatoxin corresponds to the oxygen at C_9 in phorbol.

All of the three models, A–C, reflect the fact that debromaplysiatoxin is a very effective competitive inhibitor of PDBU binding to PKC. Model A, however, suggests that debromoaplysiatoxin should bind more tightly than PDBU does, and this conflicts with the bioassay data, which show that PDBU, with a K_i value[8] of 0.000 78 μM, has a higher affinity than debromoaplysiatoxin (0.02 μM). Both model B and model C predict that while debromopalysiatoxin will compete well with PDBU, it will have a lower affinity, as is in fact the case. The difference between these two models is that model C leads to about 2.89 kcal/mol larger energy gain upon binding than does model B. From the energetic point of view, when a molecule interacts with the receptor binding sites, it will always adopt a binding mode in which the entire system can have the maximal energy gain. Therefore, out of the two binding modes, the one with the larger energy gain can be considered to be the preferred binding mode, and the binding energy of this orientation will be regarded as the binding energy with the models. As is discussed below, the PM3 method is known to deliver results of only approximate absolute value, and the substitution of methanol or formaldehyde for the true donors and acceptors will further affect the absolute values of the calculated energies. Nevertheless, these data suggest that model A is the poorest of the alternatives while model C is the best.

The binding energy of 1,2-dipalmitoyl-*sn*-glycerol (**5a**) in each of the three models A–C was determined in the same way, and the results are also shown in Table 2. In orientation α, oxygen O_2 of **5a** corresponds to the C_3 oxygen of phorbol, and oxygen O_3 of **5a** corresponds to the C_9 oxygen in phorbol. This orientation of the triglyceride was clearly preferred since the inverse orientation leads to weaker binding in all three models. It is interesting to



1	H	4.586	0.861	0.108
2	O	5.424	0.395	0.123
3	C	6.402	1.311	0.506
4	H	7.343	0.755	0.465
5	H	6.241	1.684	1.525
6	H	6.454	2.171	-0.177
7	H	3.661	3.401	0.408
8	O	3.772	4.315	0.674
9	C	4.671	4.902	-0.218
10	H	5.422	5.425	0.381
11	H	4.130	5.630	-0.832
12	H	5.167	4.172	-0.870
13	O	-3.052	-0.669	-0.409
14	C	-3.144	-1.598	-0.633
15	H	-2.865	-1.736	-1.990
16	H	-2.955	-2.808	-2.189
17	H	-1.846	-1.405	-2.234
18	H	-3.578	-1.187	-2.618
19	C5	0.640	3.421	-0.999
20	C6	-0.696	3.978	-1.368
21	C20	-0.705	4.724	-2.684
22	O3	-0.624	3.791	-3.735
23	H3	-0.572	4.298	-4.547
24	O	-0.491	5.058	-6.193
25	C	-0.602	4.406	-7.201
26	H	-0.772	3.329	-7.170
27	H	-0.536	4.871	-8.185

Figure 7. PKC receptor site, model C. For the 27 atoms of the model, the *x*, *y*, and *z* coordinates are given in angstroms.

recall that for this compound, the simple two-point geometrical fit shown in Table 1 was unable to determine which binding manner is preferred. The strongest possible two-point binding energy between **5a** and two models, A and B, is -5.66 kcal/mol, associated with the A_α binding. As can be seen from Table 2, the diglyceride **5a** binds well with both model A_α and model B_α, although model A_α was 0.26 kcal/mol better than model B_α. With model C_α, three methanols formed hydrogen bonds with the triglyceride, with a binding energy of -7.66 kcal/mol. If the system in model C_α was reminimized with the triglyceride unconstrained, the binding energy between **5a** and three methanol molecules increased only to -8.40 kcal/mol, suggesting that in the original fit, the interactions between **5a** and the three methanols of model C_α had been nearly optimal. Since relaxing of the constraints did not improve the binding greatly, it seems to indicate that the conformation of the crystal structure of **5a** was indeed similar to the conformations adopted by the 1,2-diacylglycerols when bound to the PKC receptor and, moreover, that model C_α is the best of the binding-site models. Thus the data for both debromoaplysiatoxin (4) and 1,2-dipalmitoyl-*sn*-glycerol (**5a**) suggest the superiority of model C_α over the other alternatives and this indication received further support from the synthetic compounds that were studied next.

The binding energies between the six lactones (**6**–**11**) and these three models were evaluated using the same procedure as described above, with the two orientations generated by interchanging the two carbonyl oxygens. In orientation α, the oxygen labeled as "1" was allowed to bind with the binding point previously associated with the C_3 oxygen of PDBU. In model C, the smallest energy difference between the two binding orientations was 6.61 - 5.35 = 1.26 kcal/mol, for lactone **8**. Using models A and B, for compounds such as **7**, it was not possible to distinguish the two binding orientations since they had very similar binding energies. In model C, on the other hand, the two binding modes can always be easily distinguished. For the lactones **6**–**11**, the average energy difference between the C_α and C_β modes was 1.81 kcal/mol, with the C_α orientation consistently favored over the C_β orientation. Thus, once again, model C was better in terms of its ability to distinguish the two possible binding modes and the C_α model/orientation gave better binding energies in all cases. Details of model C of the PKC receptor are given in Figure 7.

The calculated binding energies for all nine compounds tested in all the three models A–C are given in Table 2.

On the basis of the energy data, models A, B and C all clearly show the synthetic lactones (6–11) to be weaker binders to PKC than either esters of phorbol (1, R = H), debromoaplysiatoxin (4), or 1,2-dipalmitoyl-*sn*-glycerol (5a), and this is consistent with the measured binding constants. The binding energies for the lactones with model A were at least 1 kcal/mol less than that for phorbol with model A. The binding energies between the lactones and model B were at least 1.62 kcal/mol less than that between phorbol and model B, and the binding energies between the lactones and model C were at least 3.09 kcal/mol less than in the phorbol case. Among the lactones themselves, three compounds, 6, 7, and 8, had larger binding energies than the other three, although lactone 8, with the highest (negative) binding energy (−6.61 kcal/mol), was slightly less potent than 6 and 7. This may be due to the need for a third binding site involving the free hydroxyl group of the lactone and the formaldehyde in the model, a point that can be resolved by including another binding site in the model, which is done below.

The results of two-point binding, listed in Table 2, show that, as had been suggested by the data on debromoaplysiatoxin and 1,2-dipalmitoyl-*sn*-glycerol, model C is superior to the alternative models, A and B, and that model/orientation C_α gives the best results. Consequently, this model/orientation was chosen for further study.

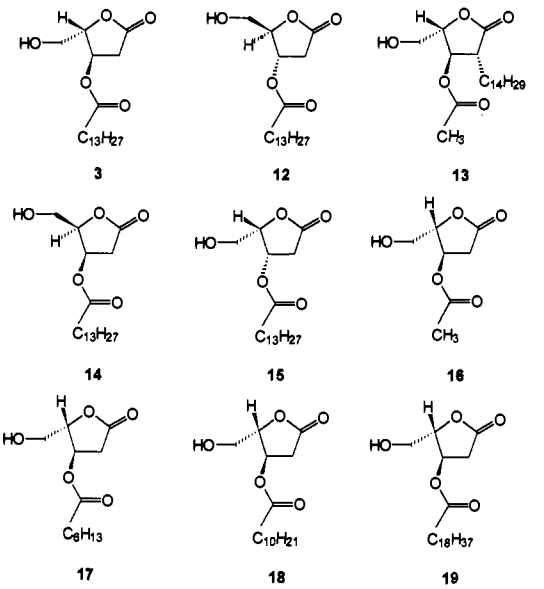
c. Inclusion of the Binding Site for the C_{20} Hydroxy Group of Phorbol: Three-Point Binding Energy. The preceding analysis of the energetics of two-point binding achieved the mapping of the pharmacophore onto the ligands by clearly establishing the orientation of each molecule in the receptor model. The calculated two-point binding energies given in Table 2 for model C can be approximately correlated with the measured binding constants, but numerous minor discrepancies remain. To account for these, the effect of the third binding site must be considered. It was assumed that the hydroxyl groups in all of the test compounds are in a position consistent with the formation of a strong hydrogen bond with the formaldehyde in the models. The three-point fits recorded in Table 1 show that the hydroxyl groups in all the compounds were close to the hydrogen-bond acceptor on PKC, thus permitting the formation of hydrogen bonds.

When calculating the binding energy at this binding site, a conformational search is necessary in order to find the hydroxyl group conformation which gives the strongest hydrogen bond with the formaldehyde in the models. This conformational search was done using Quanta, and the optimum structure was exported into PM3 where the binding energy calculations were carried out. The full (three-point) binding energies are given in Table 3, together with the measured binding constants. From the data in Table 3, it is clear that, while other factors such as entropy may play a role, a major reason for the lower binding affinities of these rigid lactones is simply their inability, for reasons of geometry, to bind strongly with the receptor site. Within the lactones, two compounds, 6 and 7, have the highest calculated binding energies and also the highest affinities (lowest K_i values).

3. Testing of the Model with Monocyclic Synthetic Lactone Analogs. Nine ribonolactones (3, 12–19) had been synthesized as part of this Laboratory's continuing effort[9] to develop synthetic PKC activators. Since they are cyclic and therefore somewhat rigid, they should not experience as large an entropy loss when binding to PKC as do the acyclic diacylglycerols.

Table 3. Experimentally Measured K_i Values (μM), PM3-Calculated Binding Energies (ΔH^{PM3}, kcal/mol), Number (N_{RB}) of Rotatable Bonds Frozen during Binding, and the log of the Estimated Aqueous Solubility (WS) for Each Ligand

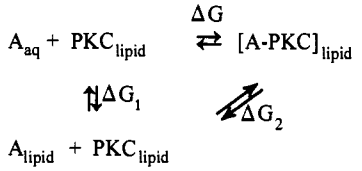
compd no.	K_i (exptl)	ΔH^{PM3}	N_{RB}	log(WS)
1 (R = COC_3H_7)	0.000 78	−12.68	2	−1.22
4	0.02	−11.35	2	−2.36
2	1.3	−10.73	7	−2.73
6	7.7	−10.02	2	−2.26
7	7.6	−9.58	2	−2.26
8	13.6	−9.07	2	−2.84
9	22.0	−8.91	2	−1.74
10	27.8	−9.05	2	−1.67
11	75.3	−8.70	4	−2.84
3	2.5	−10.44	4	−2.62
12	5.3	−10.39	4	−2.62
13	0.5	−10.44	4	−3.34
14	300	−7.49	4	−2.62
15	31	−9.02	4	−2.62
16	>1700	−10.44	4	3.62
17	187	−10.44	4	1.02
18	4.1	−10.44	4	−1.06
19	8.7	−10.44	4	−5.28



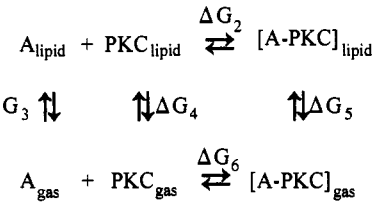
The binding energies were calculated for 3 and 12–19 in a similar way as for the bicyclic lactones by exploring different orientations and conformations in Quanta and then minimizing in PM3 the best three-point-bound conformation found. Compound 3 produced 10 conformations which were able to bind to model C. The best of these had a binding energy (enthalpy) estimated at −10.44 kcal/mol. This, with the best conformation from compound 13, which contains the identical pharmacophore, is the most effectively bound of the five lactones. Lactone 12, in its best conformation, gave a binding energy of −10.39 kcal/mol. Compound 14 gave a very low binding energy (−7.49 kcal/mol), suggesting it to be about a 150 times poorer binder than compounds 3 or 13. Finally, compound 15 gave an intermediate value for the binding energy, which is consonant with its intermediate position in terms of K_i.

The data for 3 and 12–15 are very encouraging because they suggest the model has some validity and can distinguish between ligands with similar flexibilities and similar binding affinities. The major advantage of this

Scheme 1



Scheme 2



model compared to other previously proposed pharmacophore models for PKC binding is its ability to provide a quantitative evaluation of the strength of the hydrogen bonds that the ligand can form with the binding site in PKC.

However, the data for the subgroup of compounds **3**, **16–19** show immediately that this model breaks down for these compounds, which only differ in the length of their acyl side chain—this clearly has an effect mostly on the lipophilicity of the compounds, and obviously, this is not accounted for by the present model. An additional problem is revealed by comparison of GMA (**2**) and DAT (**4**): Although their binding energies are quite similar (−10.73 and −11.35 kcal/mol, respectively), their K_i values differ by a factor of 65; here, obviously, the crucial differences are the very different flexibilities of the compounds. The model is therefore reasonable in a qualitative sense, but if it is to be refined further, these two factors must be considered in addition to enthalpy.

B. Entropy and Partitioning Effects. 1. The Relationship between Estimated Binding Energy and Measured Binding Affinities. Protein kinase C is a lipid-soluble enzyme which interacts with ligands in the lipid phase[8] and binding to it of a ligand or agonist (A) involves two steps: (1) partitioning of the agonist between the aqueous and lipid phase and (2) binding, in the lipid phase, of the agonist to PKC. These two steps are shown in Schemes 1 and 2. For the system in Scheme 1,

$$\Delta G = \Delta G_1 + \Delta G_2 \quad (1)$$

thus,

$$\Delta G_1 + \Delta G_2 = 2.303RT \log_{10} K_i \quad (2)$$

That is, the binding constant K_i is related to the free energy change (ΔG_1) associated with the partitioning of a ligand between the aqueous and lipid phases and the free energy (ΔG_2) of binding of the ligand to the PKC in the lipid phase.[8]

It is difficult to calculate free energies in a lipid environment, but the corresponding free energy changes in the gas phase can be estimated. To do this, the process of ligand binding must be further broken down into additional steps as shown in Scheme 2. From Scheme 2, it is clear that

$$\Delta G_2 = \Delta G_3 + \Delta G_4 + \Delta G_5 + \Delta G_6 \quad (3)$$

and combination of eqs 2 and 3 gives

$$2.303RT \log_{10} K_i = \Delta G = \Delta G_1 + \Delta G_3 + \Delta G_4 + \Delta G_5 + \Delta G_6 \quad (4)$$

Two quantities in eq 4, ΔG_4 and ΔG_6, pertain to the enzyme in the gas phase and are of course, unknown, and consequently, an absolute solution to eq 4 is not possible. The equation can, however, be used to establish relative binding constants. Consider two different ligands, P and Q. The foregoing consideration will lead to eqs 4P and 4Q, thus

$$2.303RT \log_{10} K_i^P = \Delta G^P = \Delta G_1^P + \Delta G_3^P + \Delta G_4^P + \Delta G_5^P + \Delta G_6^P \quad (4P)$$

and

$$2.303RT \log_{10} K_i^Q = \Delta G^Q = \Delta G_1^Q + \Delta G_3^Q + \Delta G_4^Q + \Delta G_5^Q + \Delta G_6^Q \quad (4Q)$$

If eq 4Q is subtracted from eq 4P, then

$$2.303RT(\log_{10} K_i^P - \log_{10} K_i^Q) = (\Delta G_1^P - \Delta G_1^Q) + (\Delta G_3^P - \Delta G_3^Q) + (\Delta G_4^P - \Delta G_4^Q) + (\Delta G_5^P - \Delta G_5^Q) + (\Delta G_6^P - \Delta G_6^Q) \quad (5)$$

Step 4, the equilibrium between PKC_{lipid} and PKC_{gas} involves only the enzyme, and so $\Delta G_4^P = \Delta G_4^Q$. If it is assumed that the two ligands are similar small molecules, then $\Delta G_3^P \approx \Delta G_3^Q$ and $\Delta G_6^P \approx \Delta G_6^Q$, and eq 5 reduces to

$$2.303RT(\log_{10} K_i^P - \log_{10} K_i^Q) = (\Delta G_1^P - \Delta G_1^Q) + (\Delta G_5^P - \Delta G_5^Q) \quad (6)$$

This equation shows that the difference in the binding affinities of two ligands can be estimated by computing for each the free energy change in partitioning (ΔG_1) and adding the free energy change (ΔG_5) upon binding in the gas phase (i.e., the value obtained from the modeling experiments).

The change in free energy during the partitioning process is related to the hydrophobicity of the ligand, as is discussed below. The free energy of binding (ΔG_5) can be further decomposed into enthalpy and entropy components:

$$\Delta G_5^P - \Delta G_5^Q = (\Delta H_5^P - \Delta H_5^Q) - (T\Delta S_5^P - T\Delta S_5^Q) \quad (7)$$

where we restrict ourselves in the following to entropy losses caused by immobilization of rotatable bonds of the ligand during its binding to the receptor, assuming that the translational entropy losses are approximately the same for all ligands.

Combining eq 6 and 7 gives

$$2.303RT(\log_{10} K_i^P - \log_{10} K_i^Q) = (\Delta G_1^P - \Delta G_1^Q) + (\Delta H_5^P - \Delta H_5^Q) - (T\Delta S_5^P - T\Delta S_5^Q) \quad (8)$$

Equation 8 implies that for two substrates, the one that interacts more strongly (more negative ΔH_5) with the enzyme will have the higher binding affinity (lower K_i) provided that they have similar partitioning behavior and show similar entropy changes during binding. If the substrates show marked differences in either of these regards, then the ΔH_5 values may not determine the relative K_i values. Whatever the situation, eq 8 makes it

clear that all three of these aspects must be considered in attempts to determine the binding affinity of a substrate.

If ligand Q is chosen to be a reference compound with a known K_i value, then we can rearrange eq 8 to make the dependency of K_i^P on the known (measured or modeled) quantities clearer:

$$2.303RT \log_{10} K_i^P = [(\Delta G_1^P - \Delta G_1^Q) + (\Delta H_5^P - T\Delta S_5^P)] + [2.303RT \log_{10} K_i^Q - (\Delta H_5^Q - T\Delta S_5^Q)] \quad (9)$$

The second term of the right side of this equation is a constant, which means that $\log_{10} K_i^P$ is linearly related to the sum of $(\Delta G_1^P - \Delta G_1^Q)$ and $(\Delta H_5^P - T\Delta S_5^P)$. Obtaining ΔH_5^P values had been the goal of constructing the binding-site model and calculating binding energies from it. $T\Delta S_5^P$ is the energy penalty to pay for immobilization of rotatable ligand bonds upon binding to the receptor, and this is discussed in the following section. The value of $(\Delta G_1^P - \Delta G_1^Q) = \Delta\Delta G_1$, finally, which expresses the partitioning characteristics of the compound, is related to its water solubility, and this is further analyzed in the subsequent section.

If all terms appearing in eq 9 could be calculated theoretically (or known experimentally, for the case of K_i^Q) without any uncertainty, then direct application of this equation should yield a perfect prediction of the K_i for any new compound (within the fundamental applicability limits of the model, of course). Uncertainties exist, however, for all of these terms, and therefore a maximally predictive model was obtained by multiple regressive fit of the model to the known K_i values, which is detailed below.

2. Entropic Energy Changes During Binding. Recent studies by Williams[41] have shown that entropy loss during substrate–enzyme binding is due primarily to the restrictions imposed during binding upon free translation and internal rotation of the substrate, a picture which is consonant with the accepted qualitative view of the process. Williams estimated that for each rotor restricted as a result of the substrate's binding to the enzyme, the entropy loss ($T\Delta S$ at 300 K) is between −0.38 and −0.86 kcal/mol. In a study of the binding of peptide ligands to ristocetin A,[42] the $T\Delta S$ value was estimated to be −0.89 ± 0.2 kcal/mol per rotatable bond. If one would have to settle on a specific number, without having any other information, an intermediate value on the order of −0.5 kcal/mol would probably be a natural choice. In our approach, the multiple regression delivered an average value of $T\Delta S$ that was somewhat lower than the above range of values but which seems not at odds with the accepted qualitative view of the process.

3. Partitioning of the Ligand. In order to evaluate eq 9, one needs the free energy change associated with the partitioning of the ligand between the aqueous and the lipid phases, $(\Delta G_1^P - \Delta G_1^Q) = \Delta\Delta G_1$, but, unfortunately, these quantities are not directly known. An approximate value for $\Delta\Delta G_1$ can, however, be obtained through the following consideration. Assuming that partitioning (at least in the PKC system studied here) is a passive process, then the efficiency with which a substrate transfers into the lipid phase will be determined by the hydrophobicity of the substrate. The hydrophobicity, in turn, is a function of the water solubility (WS) of the compound, which can be quite accurately estimated by the method developed by Klopman *et al.*[43] If one now takes a number of compounds that, e.g., vary only in their (lipophilic) side chains, their binding energies (ΔH_5) and entropic energy losses ($T\Delta S_5$) should all be the same, and consequently, the variation in their K_i (a function f) should reduce to a function f' of (only) the remaining variable, $\Delta\Delta G_1$, and hence of WS. We thus have achieved

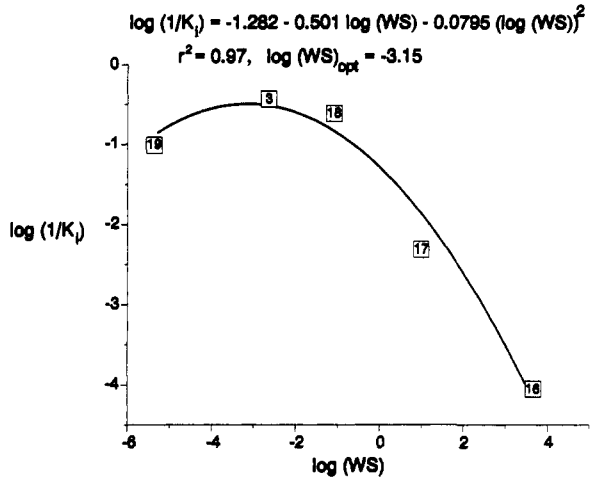


Figure 8. Plot of log(WS) vs $\log(1/K_i)$.

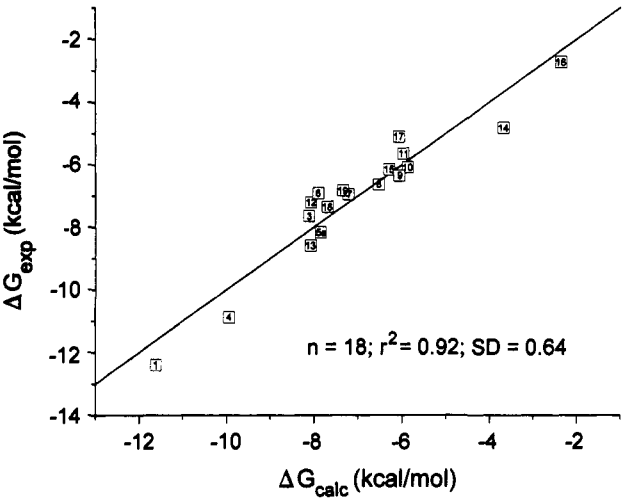


Figure 9. Plot of measured vs estimated free energy of binding.

$$\log_{10} K_i = f(\Delta\Delta G_1, \Delta H_5, T\Delta S_5) \Rightarrow f'(\Delta\Delta G_1) \quad (10)$$

and from (8) it follows

$$f'(\Delta\Delta G_1) = a\Delta\Delta G_1 + b \quad (11)$$

with two constants a and b expressing the lack of knowledge about the absolute values of ΔG_1. On the other hand, we have assumed a functional dependency

$$\log_{10} K_i = g(\mathrm{WS}). \quad (12)$$

Inversion of (10) and entering (11) in the result yields

$$a\Delta\Delta G_1 + b = \log_{10} K_i = g(\mathrm{WS}) \quad (13)$$

In order to obtain a functional relationship (i.e., the function g) between WS and K_i, the logarithm of the water solubility for five of the monocyclic lactones (**3**, **16**–**19**) that have a varying side-chain length was calculated[43] and plotted against $\log_{10}(1/K_i)$ (Figure 8). Although it is difficult with so few data points to determine the exact

Table 4. Contributing Values of Each Variable Term in Equation 16, the K_i-Derived Free Energy Changes, the Free Energy Changes Calculated from Equation 16, and the Difference between the Measured and Calculated Free Energy Changes (kcal/mol)

compd no.	$1.505\Delta H^{PM3}$	$0.235N_{RB}$	$[0.139(\log(WS) + 2.823)^2]$	ΔG_{exp}	ΔG_{calc}	$\Delta G_{calc} - \Delta G_{exp}$
1 (R = COC_3H_7)	-19.08	0.47	0.36	-12.39	-11.60	0.79
4	-17.08	0.47	0.03	-10.88	-9.92	0.96
5a	-16.15	1.75	0.001	-8.16	-7.84	0.32
6	-15.08	0.47	0.04	-6.97	-7.91	-0.94
7	-13.67	0.47	0.04	-6.98	-7.24	-0.26
8	-13.65	0.47	0.00	-6.64	-6.52	0.12
9	-13.41	0.47	0.16	-6.35	-6.12	0.23
10	-13.62	0.47	0.18	-6.21	-6.30	-0.09
11	-13.09	0.94	0.00	-5.63	-5.96	-0.33
3	-17.51	0.94	0.01	-7.64	-8.11	-0.47
12	-15.64	0.94	0.01	-7.20	-8.03	-0.83
13	-15.71	0.94	0.04	-8.59	-8.08	0.51
14	-11.27	0.94	0.01	-4.80	-3.67	1.13
15	-13.58	0.94	0.01	-6.15	-5.97	0.18
16	-15.71	0.94	5.77	-2.72	-2.35	0.37
17	-15.71	0.94	2.05	-5.09	-6.06	-0.97
18	-15.71	0.94	0.43	-7.35	-7.68	-0.33
19	-15.71	0.94	0.84	-6.90	-7.28	-0.38

mathematical equation for the relationship between aqueous solubility and binding affinity, it seems to be suggested from the data points and from fundamental considerations that compounds which are either too hydrophobic or too hydrophilic will not be able to achieve a high binding affinity; i.e., a nonlinear relationship is to be expected. Quadratic fit indeed produced a very high correlation ($r^2 = 0.97$) for the interpolation curve given in Figure 8. From the data points, the optimal water solubility ($\log(WS_{opt})$) should be close to -3.0; the regression yielded a $\log(WS_{opt})$ of -3.15.

If we choose a reference compound with (or with a value close to) the optimum water solubility and set its (reference) ΔG_1 value, ΔG_1^Q, to zero, then the constant b in (13) becomes zero also, and $\Delta\Delta G_1$ becomes a function of WS/WS_{opt} only, which in turn can be computed from the difference in the $\log_{10}(1/K_i)$ values between the reference compound and the compound of interest. With the fundamental thermodynamic relationship as expressed in (4) or (8), we thus finally obtain

$$\Delta\Delta G_1 = -2.303RT\,\Delta\log_{10}(1/K_i) \quad (14)$$

with the dependency of $\Delta\log_{10}(1/K_i)$ on WS/WS_{opt} determined by the function given in Figure 8, and from this a numerical value for $\Delta\Delta G_1$ can easily be established.

This was, however, not done directly, since the optimally predictive model was derived by multiple regression, but these considerations provided the theoretical underpinnings for using the quadratic term $c_3(\log(WS) + c_4)^2$ in the regression equation discussed below.

4. Development of the Optimum Model by Multiple Regression. Following the foregoing discussion, a quantitative model was established that correlates the K_i values with the PM3-calculated binding energy, ΔH^{PM3} (corresponding to the ΔH_5 in (7), etc.), the entropy loss incurred upon binding to PKC, and the hydrophobicity of the ligand expressed as a function of log(WS). The following regression was used:

$$2.303RT\log_{10}(K_i) = c_0 + c_1\Delta H^{PM3} + c_2N_{RB} + c_3(\log(WS) + c_4)^2 \quad (15)$$

with the five coefficients $c_0 - c_4$. As the variable accounting for the entropy loss, we simply used the number of rotatable bonds immobilized upon binding, N_{RB}, and $[c_3(\log(WS) + c_4)^2]$ is the contribution of the hydrophobicity (represented by the water solubility WS) of the ligand.

The coefficients $c_0 - c_4$ were derived by fitting eq 15 to the measured K_i values of the 18 compounds listed in Table 4, using multiple regression. This yielded

$$2.303RT\log_{10}(K_i) = 6.62 + 1.505\Delta H^{PM3} + 0.235N_{RB} + 0.139(\log(WS) + 2.832)^2 \quad (16)$$

with a very good correlation of $r^2 = 0.917$ (SD = 0.64, $F(4,13) = 35.69$). A plot of this equation is shown in Figure 9. The Student's t test indicated that all three parameters are statistically significant, the most, second, and least significant ones being ΔH^{PM3}, log(WS), and N_{RB}, respectively.

Equation 16 reveals a number of interesting points. The multiple regression produced a scaling factor for the PM3-derived binding energies of 1.505. It is known that PM3 has the tendency of underestimating the energy of a typical hydrogen bond system by about 1–2 kcal/mol when the system is in its optimal geometry. For example, for the $(H_2O)_2$ dimer, the PM3-calculated hydrogen bond energy has been shown[41] to be -3.48 kcal/mol as compared to the experimental value of -5.1 kcal/mol.[44] In other words, the calculated value would have to be increased by approximately 50% to agree with the measured value—and this is exactly the scaling factor that was delivered by the multiple regression in this study. The entropic energy loss incurred upon freezing the rotatable bonds was found to be 0.235 kcal/mol per rotatable bond. This value is somewhat lower than the range given above that was found in other studies (0.86–0.36 kcal/mol per rotatable bond), but it is still on the same order and does not seem to be unreasonable *per se*. The fact that, in our data set, the variation of N_{RB} is quite limited may have contributed both to this value and the (relatively) low significance of this parameter. The optimum log(WS) value from eq 16 is -2.832, which is consistent with the value obtained from Figure 8.

Discussion

The major advantage of this hypothetical receptor model compared to other previously proposed pharmacophore models for PKC binding is its ability to determine if a ligand can form strong hydrogen bonds with the binding site in PKC. The superiority of this model to others stems

from the fact that the hydrogen-bond interaction depends not only on the distance between the hydrogen-bond donor and hydrogen-bond acceptor but also on the geometry of the hydrogen bonds. The strength of a hydrogen bond depends on four geometrical parameters: the distance between the hydrogen bond donor atom and the hydrogen bond acceptor atom, two bond angles, and one torsion angle. Use of a pharmacophore which only contains the positions of the three essential oxygens in the phorbol ester implies the ability to control only the positions of these oxygens in the new compounds which will be expected to match the corresponding oxygens in phorbol. If however, these oxygens have improper orientations in space, they will not be able to form strong hydrogen bonds with the binding points on the PKC receptor, and as a result, the designed compound will have a low binding affinity for PKC.

A further advantage of the approach presented is that it implicitly takes care of the conformational problem encountered with these kinds of pharmacophore mapping, receptor (site) modeling, and QSAR approaches. Many other techniques, such as, e.g., CoMFA, offer no explicit provision for handling conformational flexibility and the associated energetics, although attempts at circumventing this problem have been proposed.[45] Here, an eventual energy penalty to be paid for a conformational deformation necessary for accommodating the ligand in the receptor site is automatically taken into account by the PM3 energy calculation that is performed for the entire pseudoreceptor/ligand complex—at least to the extent that energy minimization(s) and conformational search(es) can be, and are, performed that explore the entire energy surface of this complex.

By the same token, however, this represents one of the disadvantages of this approach: semiempirical energy calculations are relatively computer-intensive, and entire conformational searches obviously even more so although the situation can be, and was, in the present study, somewhat ameliorated by using a lower-level approach, such as molecular mechanics, to perform the actual conformational search. Furthermore, this approach necessitates a fair amount of human intervention, and for those reasons it would not lend itself easily to rapid screening of a large number of structures. One could, however, speculate about using a two-step approach in order to combine both tasks. The high-level calculations would be used to establish the receptor model, but then, for the screening, there would be substituted a carefully parametrized lower-level approach such a molecular mechanics that carefully models the details of the hydrogen-bond interactions. The GRID technique[24] is an example of the latter approach.

Although multiple regression was used to derive the optimal quantitative model, the parameters used in the model have a real physical meaning. Especially the binding enthalpies, as calculated by PM3 from the three-point binding model, already provide a fairly good ranking all by themselves for those compounds (4, 6, 8–12) for which the other two factors, flexibility and hydrophobicity, are not of overwhelming influence. This is consistent with the view[32] that, in many biological systems, hydrogen bonds constitute the most important interaction in ligand–macromolecule complexes, and it is also in agreement with the results of an analysis of the energetics of a series of enzyme–ligand complexes whose structures have been established by X-ray measurements.[46] This was further supported by tests conducted where one or the other of the coefficients $c_0 - c_4$ was held fixed at a certain value during the multiple regression. If c_1 was set to 1, i.e., ΔH^{PM3} was used "as is", a correlation (r^2) of still 0.724 was obtained. If, in addition to $c_1 \equiv 1$, c_2 was set to 0.5 (kcal/mol per immobilized rotatable bond), r^2 was reduced only to a value of 0.631. The correction of the PM3 energies by the scaling factor 1.5 does, however, seem to be an important step in creating an optimal model, since restricting only c_2 (to 0.5), yielded a model with a correlation that was nearly as high ($r^2 = 0.891$) as that of the fully optimized model.

Although scaling the PM3 energies is some kind of solution to the shortcomings of this semiempirical method, the preponderance of the binding enthalpy in the model derived shows that improvements in the methodology used to calculate the geometry parameters (hydrogen bond lengths, angles, and torsion angles and so on) and the energy of the hydrogen bond system could be expected to further increase the quality of the model. It has been shown[31] that PM3 is superior to other semiempirical methods, such as MNDO and AM1, in predicting the geometrical parameters for a hydrogen bond system, but it underestimates a typical hydrogen bond length by about 0.1–0.2 Å. Our PM3-derived data for the methanol–formaldehyde system also produced a low value for this bond length. The shorter hydrogen bond length predicted by PM3 is probably due to its underestimate of the core–core repulsion energy.[31] However, this error in the bond length does not present a very serious problem because it can be corrected systematically. For the quantitative receptor site model, improvements in the accuracy of the energy determination would obviously be of greater importance.

The model clearly suggests that PDBU (1, R = COC_3H_7), although being the most potent compound already, could be made even more active. With its log(WS) of −1.22, it is less than optimally hydrophobic; experimental observation[8] has shown that only about 3.1% of PDBU partitions from the aqueous phase to the lipid bilayer. These data suggest that modification of the PDBU side chains to make the compound more hydrophobic would increase its potency. Our calculations showed that addition of three or four more carbons to the side chains of the esters attached to C_{12} or C_{13} of PDBU, will lead to a log (WS) value close to the optimal value of −3.0. log(WS) equals −2.78 with three more carbons and −3.30 with four more carbons. This conclusion has been confirmed experimentally,[47] albeit by using a biological endpoint, rather than binding as the measure of activity. The quadratic dependency on log(WS) has clearly be shown, too, since attaching too long a side chain to the ester groups on C_{12} and/or C_{13} has been demonstrated[48] to diminish the compound's binding affinity by making it too hydrophobic.

The established receptor model should help offer a new approach to the design of new PKC agonists. When designing a new agonist, it is necessary to ensure not only that the three oxygens in the compound are in the proper locations but also that they are appropriately oriented. It is not trivial to design a readily synthesized structure which satisfies these conditions of location and orientation, but it should be possible to evaluate proposed synthetic targets in terms of their ability to bind to this receptor model. If a proposed structure is determined to have reasonably high binding energies with the model, as well as an apposite

geometry, it then has to be further investigated whether too high an entropy loss and/or an inapropriate hydrophobicity might still cause an untolerable decrease in its binding affinity with the PKC receptor, and only if this is not the case should the compound be included in the synthetic plan.

However, as with all modeling approaches, this model's principal limitations should not be overlooked. The model that has been established here is based upon our knowledge of the binding site of PKC. It contains no information about other atoms of PKC. However, for a regulatory agonist to bind effectively with this enzyme, not only are the positions and orientations of the three pharmacophore oxygens important but it clearly is necessary that the ligand fit within the volume available to it in the binding cavity. Little information is available on the dimensions of this cavity, but it may be possible to discover its boundaries by modification of known ligands. In the absence of such knowledge, however, the calculated binding energy should be treated as the possible maximal binding energy. This will be especially true if the molecule has a bulky group near one of the binding sites. The bulky group may prevent the binding of the molecule to the receptor. In our quantitative model, the role played by a ligand side chain in the interaction with the PKC receptor was accounted for by use the log(WS) term which is a surrogate for the overall lipophilicity of the ligand. It appears, however, that the shape and the orientation of this side chain are important in the binding process. This kind of problem can possibly be addressed by using the active analog approach[49] or the CoMFA methodology.[22,23]

Our understanding of the entropy changes in the binding process are still very rudimentary, and the estimates made here of the entropy losses during PKC binding are based on limited information. It seems, for example, that the "one-size-fits-all" procedure of assigning *one* $T\Delta S$ value to all different kinds of rotatable bonds irrespective of a possibly hindered rotation and the consequently different amount of phase space (states) available to each rotor is a rather simplistic approach. In the future, a fuller understanding of the entropy changes in the binding process will support more accurate estimation of these data.

Conclusion

A model of the receptor site in protein kinase C has been developed by examining the interactions between probe molecules and crucial binding points of a rigid substrate with high binding affinity for the enzyme. The approach proposed, receptor mapping, uses probe molecules to construct a hypothetical pseudoreceptor. It allows construction of the receptor model by detailed examination of the hydrogen-bond interactions between probe molecules and a receptor template. This structural model, by itself, shows already a good correlation with the measured binding affinities. The quantitative model can, however, substantially be improved by including the effect of entropic energy loss upon binding of flexible ligands to the enzyme as well as that of hydrophobicities that are markedly different from the value that is optimal for binding to PKC. The success of this approach relies on the knowledge of the pharmacophore in PKC activators, whose conformation when bound to the enzyme must be deduced, but the choice and the quality of the computational method are also important.

This approach is particularly useful at our current stage of knowledge since the structures of most of the receptors of interest to the medicinal chemist have not yet been determined by X-ray crystallography or by NMR. The approach outlined here can provide useful information about the receptor site in the absence of such direct measurements. It is not limited to PKC but should also be applicable to other enzyme–ligand systems.

References

(1) Nishizuka, Y. The Role of Protein Kinase C in Cell Surface Signal Transduction and Tumor Promotion. *Nature* **1984**, *308*, 693–698.
(2) Berridge, M. J. Inositol Triphosphate and Diacylglycerol: Two Interacting Second Messengers. *Annu. Rev. Biochem.* **1987**, *56*, 159–193.
(3) Nishizuka, Y. Studies and Perspectives of Protein Kinase C. *Science* **1986**, *233*, 305–312.
(4) Nishizuka, Y. The Molecular Heterogeneity of Protein Kinase C and its Implications for Cellular Regulation. *Nature* **1988**, *334*, 661–668.
(5) Cantley, L.; Fleischman, L.; Whitman, M. The role of lipid-derived second messengers in cell growth and transformation. *Anticancer Drug Design* **1987**, *2*, 129–138.
(6) Rahmsdorf, H. J.; Herrlich, P. Regulation of Gene Expressors by Tumor Promoters. *Pharm. Ther.* **1990**, *48*, 157–188.
(7) Blumberg, P. M. Protein Kinase C as the Receptor for the Phorbol Ester Tumor Promoters: Sixth Rhoads Memorial Award Lecture. *Cancer Res.* **1988**, *48*, 1–8.
(8) Sharkey, N. A.; Blumberg, P. M. Comparison of the Activity of Phorbol-12-myristate-13-acetate and the Diglyceride Glycerol-1-Myristate-2-Acetate. *Carcinogenesis* **1986**, *7*, 677–679.
(9) Teng, K.; Marquez, V. E.; Milne, G. W. A.; Barchi, J. J.; Kazanietz, M. G.; Lewin, N. E.; Blumberg, P. M.; Abushanab, E. A. Conformationally Constrained Analogs of Diacylglycerol. Interaction of γ-Lactones with the Phorbol Ester Receptor of Protein Kinase C. *J. Am. Chem. Soc.* **1992**, *114*, 1059–1070.
(10) Rando, R. R.; Kishi, Y. Structural Basis of Protein Kinase C Activation by Diacylglycerols and Tumor Promoters. *Biochemistry* **1992**, *31*, 2211–2218.
(11) Wender, P. A.; Koehler, K. F.; Sharkey, N. A.; Dell'Aquila, M. L.; Blumberg, P. M. Analysis of the Phorbol Ester Pharmacophore on Protein Kinase C as a Guide to the Rational Design of New Classes of Analogs. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 4214–4218.
(12) Jeffrey, A. M.; Liskamp, R. M. Computer-Assisted Molecular Modeling of Tumor Promotors: Rationale for the Activity of Phorbol Esters, Teleocidin B, and debromoaplysiatoxin. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 241–245.
(13) Brockerhoff, H. Membrane Protein-Lipid Hydrogen Bonding: Evidence from Protein Kinase C, Diglyceride, and Tumor Promoters. *FEBS Lett.* **1986**, *201*, 1–4.
(14) Wender, P. A.; Cribbs, C. M.; Koehler, K. F.; Sharkey, N. A.; Herald, C. L.; Kamano, Y.; Pettit, G. R.; Blumberg, P. M. Modeling of the Bryostatins to the Phorbol Ester Pharmacophore on Protein Kinase C. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *8*, 7197–7201.
(15) Itai, A.; Kato, Y.; Tomioka, N.; Iitaka, Y.; Endo, Y.; Hasegawa, M.; Shudo, K.; Fujiki, H.; Sakai, S. I. A Receptor Model for Tumor Promotors: Rational Superposition of Teleocidins and Phorbol Esters. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 3688–3692.
(16) Nakamura, H.; Kishi, Y.; Pajares, M. A.; Rando, R. R. Structural Basis of Protein Kinase C Activation by Tumor Promotors. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 9672–9676.
(17) Leli, U.; Hauser, G.; Froimowitz, M. Requirements for the Activation of Protein Kinase C: Comparison of the Molecular Geometries of Phorbol and Diacylglycerol. *Mol. Pharmacol.* **1990**, *37*, 286–295.
(18) Doweyko, A. M. The Hypothetical Active Site Lattice. An Approach to Modelling Active Sites from Data on Inhibitor Molecules. *J. Med. Chem.* **1988**, *31*, 1396–1406.
(19) Ghose, A. K.; Crippen, G. M. Use of Physicochemical Parameters in Distance Geometry and Related Three-Dimensional Quantitative Structure-Activity Relationships: A Demonstration Using *Escherichia coli* Dihydrofolate Reductase Inhibitors. *J. Med. Chem.* **1985**, *128*, 333–346.
(20) Boulu, L. G.; Crippen, G. M. Voronoi Binding Site Models: Calculation of Binding Models and Influence of Drug Binding Data. *J. Comput. Chem.* **1989**, *10*, 673–682.
(21) Boulu, L.; Crippen, G. M. Voronoi Receptor Site Models. *Prog. Clin. Biol. Res.* **1989**, *289*, 267–277.
(22) Cramer, R. D., III; Patterson, D. E.; Bunce, J. D. Comparative Molecular Field Analysis (CoMFA). 1. Effect of Shape on Binding of Steroids to Carrier Proteins. *J. Am. Chem. Soc.* **1988**, *110*, 5959–5967.
(23) Depriest, S. A.; Mayer, D.; Naylor, C. B.; Marshall, G. R. 3D-QSAR of Angiotensin-Converting Enzyme and Thermolysin Inhibitors: A Comparison of CoMFA Models Based on Deduced and Experimentally Determined Active Site Geometries. *J. Am. Chem. Soc.* **1993**, *115*, 5372–5384.

(24) Goodford, P. J. A Computational Procedure for Determining Energetically Favorable Binding Sites on Biologically Important Macromolecules. *J. Med. Chem.* **1985**, *28*, 849–857.
(25) Kato, Y., Itai, A., Iitaka, Y. A Novel Method for Superimposing Molecules and Receptor Mapping. *Tetrahedron* **1987**, *43*, 5229–5236.
(26) Silicon Graphics Computer Systems, Mountain View, CA 94039–7311.
(27) The Quanta Molecular Modeling System is supplied by Molecular Simulations Inc., 16 New England Executive Park, MA 01803-5297.
(28) Stewart, J. J. P. Optimization of Parameters for Semiempirical Methods. I. Method. *J. Comput. Chem.* **1989**, *10*, 209–220.
(29) Stewart, J. J. P. Optimization of Parameters for Semiempirical Methods. II. Applications. *J. Comput. Chem.* **1989**, *10*, 221–264.
(30) Stewart, J. J. P. *MOPAC Manual*, 6th ed; Frank J. Seiler Research Laboratory, United States Air Force Academy: Boulder, CO 80840.
(31) Jurema, M. W.; Shields, G. C. Ability of the PM3 Quantum-Mechanical Method to Model *Inter*molecular Hydrogen Bonding Between Neutral Molecules. *J. Comp. Chem.* **1993**, *14*, 89–104.
(32) For a summary of these data, see: Jeffrey, G. A.; Saenger, W. *Hydrogen Bonding in Biological Structures*; Springer Verlag: Heidelberg, 1991.
(33) Brandl, V. F.; Röhrl, M.; Zechmeister, K.; Hoppe, W. Röntgenstrukturanalysen von Neophorbol, $C_{31}H_{39}O_9Br$, und Phorbol, $C_{20}H_{28}O_6$. *Acta Crystallogr.* 1971, *B27*, 1718–1730. The atomic coordinates of phorbol are stored in the Cambridge Structural Database under the reference code PORBET10.
(34) Edwards, M. C.; Taylor, S. E.; Willamson, E. M.; Evans, F. J. New Phorbol and Deoxyphorbol Esters: Isolation and Relative Potencies in Inducing Platelet Aggregation and Erythema of Skin. *Acta Pharm. Tox.* **1983**, *53*, 177–187.
(35) Fürstenberger, G.; Hecker, E. On the Active Principles of the Spurge Family (Euphorbiaceae) XI. [1] The Skin Irritant and Tumor Promoting Diterpene Esters of *Euphorbia tirucalli* L. Originating from South Africa. *Z. Naturforsch.* **1985**, *40C*, 631–646.
(36) Moore, R. E.; Blackman, A. J.; Cheuk, C. E.; Mynderse, J. S.; Matsumoto, G. K.; Clardy, J.; Woodard, R. W.; Craig, J. C. Absolute Stereochemistries of the Debromoaplysiatoxins and Oscillatoxin A. *J. Org. Chem.* **1984**, *49*, 2484–2489.
(37) Smith, S. O.; Kustanovich, I.; Bhamidipati, S.; Salmon, A.; Hamilton, J. A. Interfacial Conformation of Dipalmitoylglycerol and Dipalmitoylphosphatidylcholine in Phospholipid Bilayers. *Biochemistry* **1992**, *31*, 11660–11664.
(38) Goto, M.; Kodali, D. R.; Small, D, M.; Honda, K.; Kozawa, K.; Uchida, T. Single Crystal Structure of a Mixed-Chain Triacylglycerol: 1,2-Dipalmitoyl-3-acetyl-sn-glycerol. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 8083–8086.
(39) (a) Lee, J.; Teng, K.; Marquez, V. E. Synthesis of Two Diaglycerol Analogues Having a bis-Butyrolactone Skeleton. *Tetrahedron Lett.* **1992**, *33*, 1539–1542. (b) Lee, J.; Marquez, V. E.; Lewin, N. E.; Blumberg, P. M. Synthesis of Two Rigid Diacylglycerol Analogues Having a Perhydro Furo-[3,4-b]Furan bis-γ-Butyrolactone Skeleton. 2. *Tetrahedron Lett.* **1993**, *34*, 4313–4316. (c) Lee, J.; Marquez, V. E.; Bahador, A.; Kazanietz, M. G.; Blumberg, P. M. Synthesis of Two Rigid Diacylglycerol Analogues Having a Perhydro Furo-[3,2-b]Furan bis-γ-Butyrolactone Skeleton. 3. *Tetrahedron Lett.* **1993**, *34*, 4317–4320.
(40) The enzymological data cited here were measured at 310 K, while the chemical calculations were carried out on the assumption of an ambient temperature of 298 K. No attempt was made to deal with this discrepancy because it is not expected to alter any rank orders that are derived and is known to have little effect even upon absolute values of, for example, enzyme affinity constants.
(41) Searle, M. S.; Williams, D. H.; Gerhard, U. Partitioning of Free Energy Contributions in the Estimation of Binding Constants: Residual Motions and Consequence for Amide-Amide Hydrogen Bond Strengths, *J. Am. Chem. Soc.* **1992**, *114*, 10697–10704. Searle, M. S.; Williams, D. H. The Cost of Conformational Order: Entropy Changes in Molecular Associations, *J. Am. Chem. Soc.* **1992**, *114*, 10690–10697.
(42) Gerhard, U.; Searle, M. S.; Williams, D. H. The Energy Change of Restricting a Bond Rotating in the Binding of Peptide Analogues to Vancomycin Group Antibiotics. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 803–808.
(43) Klopman, G.; Wang, S.; Balthasar, D. M. Estimation of Aqueous Solubility of Organic Molecules by the Group Contribution Approach. Application to the Study of Biodegradation. *J. Chem. Inf. Comput. Sci.* **1992**, *32*, 474–482.
(44) Dill, J. D.; Allen, L. C.; Topp, W. C.; Pople, J. A. A Systematic Study of the Nine Hydrogen-Bonded Dimers Involving NH_3, OH_2 and HF. *J. Am. Chem. Soc.* **1975**, *97*, 7220–7226.
(45) Nicklaus, M. C.; Milne, G. W. A.; Burke, T. R., Jr. QSAR of Conformationally Flexible Molecules: Comparative Molecular Field Analysis of Protein-Tyrosine Kinase Inhibitors. *J. Comput.-Aided Mol. Des.* **1992**, *6*, 487–504.
(46) Unpublished work in this laboratory.
(47) Kubinyi, H. Quantitative Structure-Activity Relationships. *Arzneim.-Forsch.* **1976**, *26*, 1991–1997.
(48) Sharkey, N. A.; Blumberg, P. M. Highly Lipophilic Phorbol Esters as Inhibitors of Specific [^{3}H]Phorbol-12,13-Dibutyrate Binding. *Cancer Res.* **1985**, *45*, 19–24.
(49) Motoc, I.; Marshall, G. R. Molecular Shape Descriptors. 2. Quantitative Structure Activity Relationships Based upon Three-Dimensional Molecular Shape Descriptors. *Z. Naturforsch.* **1985**, *40a*, 1114–1120. Motoc, I.; Marshall, G. R.; Labanowski, J. Molecular Shape Descriptors. 3. Steric Mapping of Biological Receptors. *Z. Naturforsch.* **1986**, *40a*, 1121–1127.