MOLECULAR PHARMACOLOGY, 22:11-19

Distance Geometry Analysis of the Benzodiazepine Binding Site

GORDON M. CRIPPEN

Department of Chemistry, Texas A&M University, College Station, Texas 77843

Received December 28, 1981; Accepted March 11, 1982

SUMMARY

The *in vitro* binding constants to brain benzodiazepine receptor are known for a variety of benzodiazepines and also for four other different classes of strongly binding compounds. The binding data for a total of 29 drugs selected from these 5 classes was used to deduce a possible binding site model consisting of 15 site points and only 5 adjustable energy parameters. Even though some of the chemical structures differed radically, it was possible to fit the experimental data with a root mean square deviation of 1.1 kcal/mole. Apparently 5 non-hydrogen atoms in each ligand can occupy corresponding points in the site, and thus consitute a possible benzodiazepine pharmacophore.

INTRODUCTION

This is the fourth paper in a series on the distance geometry method for deducing the geometry and chemical nature of a macromolecular binding site, given only the chemical structures and observed free energies of binding for a series of ligands. The first paper (1) discussed the basic simplifying assumptions employed and the format for framing one's hypothesis about the modes of binding of the ligands. The second article (2) elaborated on additional algorithms deisgned to make the construction of a model site more automatic and less laborious. The third paper (3) described how stereospecific binding was included. However, in each of these studies, the set of drugs considered was always a series of closely related analogues, even though the methodology could deal with structurally dissimilar compounds. The anxiolytic drugs offer an interesting test of this feature. Although it had seemed for many years that only a benzodiazepine derivative could bind to the receptor used by diazepam, recent studies have shown that a few entirely different compounds have strong affinity for the same receptor. In this paper, I deal with two questions: what are the common features of these very different drugs, and what can one deduce about the geometry and energetics of the binding site?

INPUT DATA

The distance geometry analysis starts with the experimentally determined free energies of binding of a series of ligands to the same, unique binding site. The work of Braestrup and Squires (4) apparently fulfills these assumptions, as they report K_i values for the specific binding of a series of drugs measured by the displacement of [^{3}H]diazepam bound to a rat brain preparation. The recent work of Nielsen and Braestrup (5) suggests that there may be more than one brain benzodiazepine receptor, but it is not yet clear how this might affect the observed binding constants. In any event, the conclusions of the present study do not depend sensitively on the precise values of the binding constants. The free energies of binding of DIA,[1] LOR, MED, DEM, OXA, CLO, NIT, FLU, EST, CHL, U35, U31, RIP, KC4, BRO, LIB, R48,

This work was supported by National Institutes of Health, National Institute of Arthritis, Metabolism, and Digestive Diseases Grant 1-r01-AM28140-01.

[1] The abbreviations used are: DIA, diazepam (7-chloro-1,3-dihydro-1-methyl-5-phenyl-2*H*-1,4-benzodiazepin-2-one); ZOP, zopiclone ([6-(5-chloro-2-pyridyl)-6,7-dihydro-7-oxo-5*H*-pyrolo[3,4-b]pyrazin-5-yl]4-methyl-1-piperazinecarboxylate); LOR, lorazepam; MED, medazepam (7-chloro-2,3-dihydro-1-methyl-5-phenyl-1*H*-1,4-benzodiazepine); DEM, demethyldiazepam; OXA, oxazepam; CLO, clonazepam; NIT, nitrazepam (7-nitro-1,3-dihydro-5-phenyl-2*H*-1,4-benzodiazepin-2-one); FLU, flunitrazepam (5-(2-fluorophenyl)-1,3-dihydro-1-methyl-7-nitro-2*H*-1,4-benzodiazepin-2-one); EST, estazolam (8-chloro-6-phenyl-4*H*-*s*-triazolo[4,3-*a*][1,4]benzodiazepine; CHL, chlorazepate; U35, U-35.005 (Upjohn Company, Kalamazoo, Mich.) (6-(o-chlorophenyl)-1-methyl-4*H*-*s*-triazolo[4,3-*a*][1,4]benzodiazepine); U31,U-31.957 (Upjohn Company) (1-methyl-6-phenyl-4*H*-*s*-triazolo[4,3-2a][1,4]benzodiazepine; RIP, ripazepam (1-ethyl-4,6-dihydro-3-methyl-8-phenyl-pyrazolo[4,3-*e*][1,4]diazepin-5(1*H*)-one; CCE, β-carboline-3-carboxylic acid ethyl ester (3-carboethoxy-9*H*-pyrido[3,4-b]indole); MCC, 1-methyl-β-carboline-3-carboxylic acid ethyl ester; KC4, KC-4-2846 (Kali-Chemie) (7-bromo-1-methyl-2-carbamoylmethylen-5-(2′-fluorophenyl)-1,3-dihydro-2*H*-1,4-benzodiazepine); BRO, bromazepam (7-bromo-1,3-dihydro-5-(2-pyridyl)-2*H*-1,4-benzodiazepine-2-one; LIB, librium (chlordiazepoxide); R48, Ro5-4864 (4′-chlorodiazepam); ZOX, chlorozoxazone (5-chloro-2(3*H*)-benzoxazolone); PRO, chlorpromazine (2-chloro-*N*,*N*-dimethyl-10*H*-phenothiazine-10-propanamine); THO, tetrahydro-β-carboline-3-carboxylic acid; THM, tetrahydro-β-carboline-3-carboxylate methyl ester; THE, tetrahydro-β-carboline-3-carboxylate ethyl ester; TMO, 1-methyl-tetrahydro-β-carboline-3-carboxylic acid; TME, 1-methyl-tetrahydro-β-carboline-3-carboxylate ethyl ester; CHO, β-carboline-3-carboxylic acid; and CHM, β-carboline-3-carboxylate methyl ester.

0026-895X/82/040011-09$02.00/0

ZOX, and PRO were calculated from ref. 4. Other compounds covered by the same reference were excluded either because of doubt concerning their chemical structure (U 39.219, Ro 5-3027, Ro 5-3590, Ro 5-2904, Ro 5-4528, Ro 5-5807, Ro 5-3785, and Ro 5-3636) or because of excessive conformational freedom (flurazepam and serum albumin). ZOP binding data were determined by Blanchard *et al.* (6), using a very similar *in vitro* assay. The binding data for CCE, MCC, THO, THM, THE, TMO, TME, CHO, and CHM were also determined by essentially the same method (7). A good deal more is known about the structural requirements for activity from many other experiments (e.g., ref. 8), but since these were not based on the same sort of assay, they were not included in this study. Table 1 includes the full list of the 29 compounds I have treated, along with their observed binding energies. A diligent search of the literature through 1980 failed to locate more compounds measured by the same assay, having published chemical structure, with fewer than eight rotatable bonds.

The next step in the analysis is to choose a representation of the ligand molecules. Initially each ligand was described in terms of the Cartesian coordinates of every atom, including hydrogen atoms. Coordinates were obtained from published crystal structures, but although there are several such studies in the literature for benzodiazepines, they are lacking for most of the 18 compounds under consideration. However, Gilli and co-workers (9) have shown that the benzodiazepine ring system varies only slightly from derivative to derivative, so I took the diazepam coordinates of Camerman and Camerman (10) as the basis for the benzodiazepines of unknown crystal structure. Standard fragments from other crystal structures were joined to this to form the other benzodiazepine analogues using a newly developed computer program. Coordinates for LOR (11), MED (12), LIB (13), ZOX (14), and PRO (15) were taken directly from their respective crystal structures. In a similar fashion, ZOP, CCE, MCC, THO, THM, THE, TMO, TME, CHO, and CHM were constructed from fragments located by searches of the Cambridge Crystallographic Data tapes.

The all-atom coordinates of each ligand were converted in the usual way into upper and lower bound

TABLE 1

Binding to the benzodiazepine receptor

Ligand[a]	Final proposed binding mode[b]									ΔG[c]	
	s1	s2	s3	s4	s5	s6	s7	s8	s9	Observed	Calculated
										kcal/mole	
DIA	N1	C14	C18	C8	N4	O2	0	C11	C9	−11.0[d]	−10.2
ZOP	0	0	0	0	0	N16	Cl14	C12	N7	−10.7[e]	−9.2
LOR	N2	C22	C30	C17	N9	O3	Cl22	C12	C19	−11.6[d]	−11.7
MED	N5	C25	C33	C20	N12	0	0	C15	C22	−7.4[d]	−8.1
DEM	N2	C20	C28	C14	N8	O4	0	C10	C16	−11.0[d]	−10.2
OXA	N2	C19	C27	C13	N7	O4	0	C9	C15	−10.0[d]	−10.2
CLO	N2	C20	C27	C14	N8	O4	Cl32	C10	C16	−11.9[d]	−11.7
NIT	N2	C20	C28	C14	N8	O4	0	C10	C16	−10.5[d]	−10.2
FLU	N1	C19	C26	C13	N7	O3	F35	C9	C15	−11.7[d]	−13.0
EST	N3	C22	C30	C16	N10	N5	0	C12	C18	−10.8[d]	−10.1
CHL	N2	C19	C27	C13	N7	O4	0	C9	C15	−10.3[d]	−10.2
U35	N2	C22	C29	C16	N9	N4	Cl35	C11	C18	−11.8[d]	−11.6
U31	N2	C22	C30	C16	N9	N4	0	C11	C18	−9.8[d]	−10.1
RIP	0	C16	C24	0	N8	0	0	C10	N12	−9.2[d]	−8.8
CCE	0	N19	0	C27	N3	0	0	C5	0	−11.5[f]	−9.5
MCC	0	C26	0	C1	0	0	0	C4	N2	−7.7[f]	−8.0
KC4	N1	C19	C26	C13	N7	0	F41	C9	C15	−12.1[d]	−10.9
BRO	N2	N20	C27	C14	N8	O4	0	C10	C16	−10.3[d]	−12.2
LIB	N2	C26	C34	C20	N13	N4	0	C16	C22	−8.5[d]	−10.1
R48	0	0	0	0	0	0	Cl32	C15	N1	−5.4[d]	−7.2
ZOX	O4	0	0	C14	0	0	0	0	N2	−5.4[d]	−5.8
PRO	0	N9	0	0	0	0	0	C2	0	−5.5[d]	−5.8
THO	0	0	0	C19	0	N4	0	C12	0	−5.9[f]	−6.5
THM	O26	0	0	0	0	0	0	C28	N4	−6.9[f]	−7.3
THE	O26	0	0	C31	N4	0	0	C8	0	−7.6[f]	−8.3
TMO	0	0	0	C18	0	N3	0	C11	0	−5.0[f]	−6.5
TME	0	C34	0	0	0	0	0	C27	N3	−5.0[f]	−6.6
CHO	0	N19	0	0	0	0	0	C5	0	−6.5[f]	−5.8
CHM	0	N19	0	0	N3	0	0	C5	0	−10.9[f]	−8.0

[a] For abbreviations of the ligand names, see Footnote 1.

[b] Binding of atoms, labeled as in Fig. 1, to site points, s1–s9. Since s10–s15 are energetically repulsive to all atom types, there is no binding to them. A zero denotes a vacant site point.

[c] Gibbs' free energies of binding at 298 K.

[d] Ref. 4.

[e] Ref. 6.

[f] Ref. 7.

interatomic distance matrices by sampling all possible combinations of dihedral angles about all rotatable bonds. The benzodiazepine ring system was held rigid during this process, in accordance with the crystallographic evidence. The resultant distance-bound matrices involved too many atoms for the subsequent combinatorial calculations, so all non-hydroxyl hydrogen atoms were removed, nearly all nitrogen and oxygen atoms were kept, and some carbon atoms were deleted so that usually alternate non-hydrogen atoms remained. Thus each molecule was represented by the labeled atoms in Fig. 1.

MOLECULAR DATA DECOMPOSITION

The first question to ask with such a structurally diverse set of strongly binding ligands is what geometric and chemical features do they have in common, if any? The systematic examination of molecular commonalities and differences can be highly suggestive of the binding site necessary to explain the experimental data. For this initial survey of the data set, only the strongly binding ligands are of interest since the weakly interacting ones may suffer from some structural incompatibility with the site. Consequently, I considered only the first 18 drugs in Table 1, DIA through BRO, which includes only one poorly binding ligand, MCC. The reason for MCC's weak binding must be the methyl group, C30. Since MCC includes all the structure of the strongly binding CCE, using MCC in this calculation can do no harm. (After the commonality calculation had shown promising results, I added the remaining 11 ligands to the data set for the subsequent site calculations.) The algorithm is the same as in my earlier works (2), where it was referred to as the decomposition algorithm in the "fixed mode analysis." Basically the procedure is a combinatorial search for the largest number of ligand points (labeled atoms in Fig. 1) common to all ligands such that point types match and corresponding interpoint distance ranges overlap. For the purposes of this work, I have taken the point types to be simply hydrogen, carbon, oxygen, nitrogen, fluorine, chlorine, and bromine. I have not attempted to distinguish between aliphatic and aromatic carbon atoms, for example. As far as the computer program is concerned, then, matching up C12 of DIA with C25 of MED is just as good as matching C12 of DIA with C2 of MED. Of course, if simultaneously N1 of DIA matches N5 of MED, then the distance range for C12–N1 of DIA agrees with that of C2–N5 of MED, but disagrees with that of C25–N5 of MED. In other words, the requirement that corresponding interpoint distance ranges overlap tends to force chemically reasonable correspondences between ligands in spite of considerable ligand point type ambiguity. Now a small error, δd, is tolerated in the overlap of the corresponding distance ranges. If the lower bound of the first range minus δd is less than the upper bound of the second range, and if δd plus the upper bound of the first range is greater than the lower bound of the second, then the ranges are said to overlap satisfactorily. When $\delta d = 1.5$ Å, there are 5 atoms common to the 18 ligands, as illustrated in Fig. 2. The calculation demonstrates that there are no more than five atoms in such a correspondence, but there could be alternative choices of the five atoms in some of the ligands that would equally well maintain the correspondence. Unfortunately, the algorithm itself is restricted to giving only the first match. Smaller values of δd result in finding fewer than three common atoms throughout the data set, and of course larger values yield large numbers of common atoms. Setting δd to 1.5 Å is certainly not unreasonable, given that the ligands are represented by only alternate non-hydrogen atoms (see Fig. 1), and the bond lengths are approximately 1.5 Å. Quite possibly another choice of which atoms to delete from each ligand would yield a better correspondence, but it is not feasible to explore systematically all of the possible choices.

Discovering that the first 18 ligands can be decomposed into a "base group" of 5 atoms plus a number of substituents suggests that the base group is important for binding to the benzodiazepine receptor, and that each ligand might bind in a mode that puts its base group atoms in the same places. Hence, I proposed the binding roles of Table 2 for site points s1, s2, s4, s5, and s9. Consequently I proposed in Table 1 that DIA atom N1 binds at s1, etc., and similarly for the other benzodiazepines. However, the proposed binding modes of ZOP, RIP, and CCE differ from those suggested by Fig. 2 for reasons discussed later in connection with refinement of the energy parameters.

The decomposition shows the common base group and then accounts for the differences in ligand structure in terms of substituent groups. For example, carbon 2′ in the phenyl ring of the benzodiazepines is accounted for in the base group, but carbon 6′ is also always represented, except for BRO, where position 6′ contains a nitrogen atom. Thus site point s3 is intended to bind these ligand points. Another frequently occurring substituent is the carbonyl oxygen atom at position 2 on the benzodiazepine ring system, for which s6 is intended. Site point s7 is for (halogen, usually) substitution at position 2′, which is thought to be important. The reasoning behind s8 is as follows. The correspondence of CCE and MCC with DIA suggested by Fig. 2 would have the sterically disallowed methyl group of MCC occupying a region near that occupied by C31 of U35. Now C30 of MCC must be a steric liability (since otherwise it would bind at least as well as CCE does), whereas C31 of U35 is not. Therefore the most favorable orientation for CCE in the site must be something other than that indicated in Fig. 2. The orientation I have chosen brings DIA C14, N4, and C11 into correspondence with CCE N19, N3, and C5, respectively, as shown in Table 1. The ring of CCE and MCC must be positioned unambiguously so that C30 of MCC is forced to strike a sterically disallowed region of the site. For that there is a geometric requirement of three binding points: s2 and s5 as already mentioned, and then s8 for C5 of CCE and the bridgehead carbon atom (C11 of DIA) for the benzodiazepines. By having several site points in the correct spatial relation to bind several of the most common ligand features simultaneously, it is relatively easy to adjust the energy parameters so that the energetically optimal binding modes are also the desired, uniform modes of Table 1. This is especially important when there are so many ligand points of the same type in one ligand. If there were few site points to

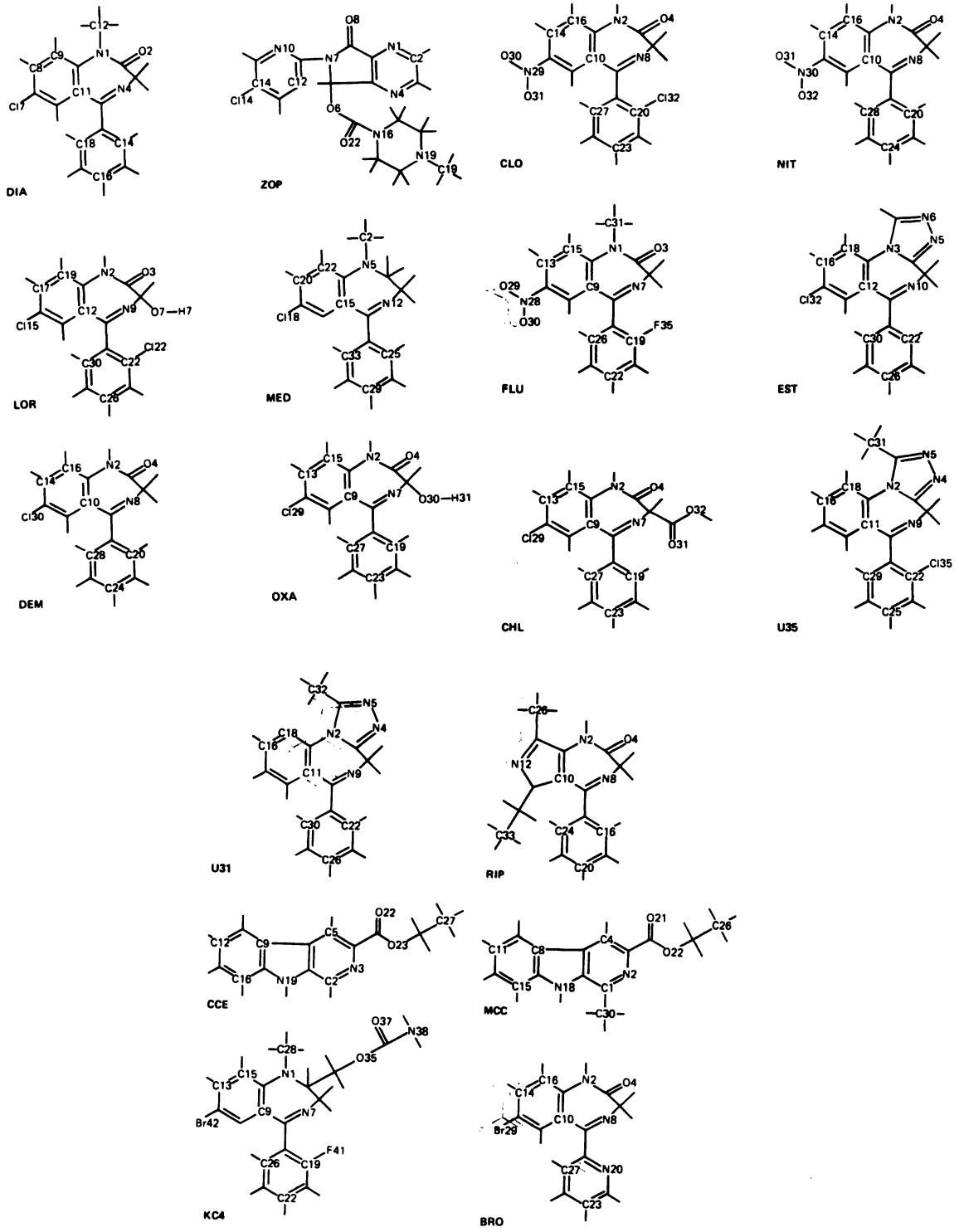

FIG. 1. *Chemical structures of the ligands*

Labels are given for those atoms included in the final representation of each molecule. All unlabeled monovalent atoms are hydrogen atoms, and all unlabeled polyvalent atoms are carbon atoms. See text for further explanation.

accommodate the ligand, there might be many different modes with similar calculated energies.

While site points s1–s9 are intended to bind the benzodiazepine derivatives in the same orientation in the site, s10 is intended to cause a change in binding mode. One of the largest changes in binding energy for the smallest alteration in structure is adding a methyl group to CCE to form MCC and worsening the binding by 3.8 kcal. As explained above, the simplest explanation is that, however CCE may lies in the site, MCC cannot do the same because of steric hindrance at its C30 methyl group. Site point s10 is therefore a "filled" site point, repulsive to all ligand point types, located so as to block MCC from enjoying the same favorable binding mode

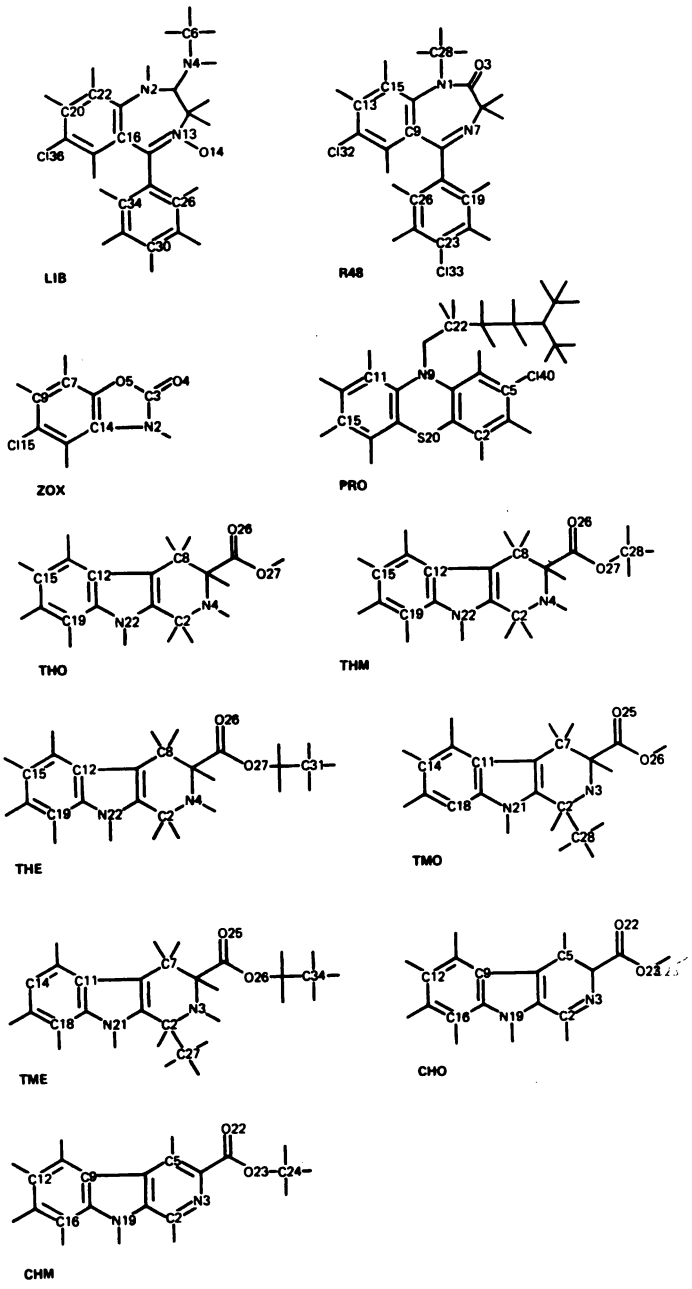

FIG. 1.—*Continued*

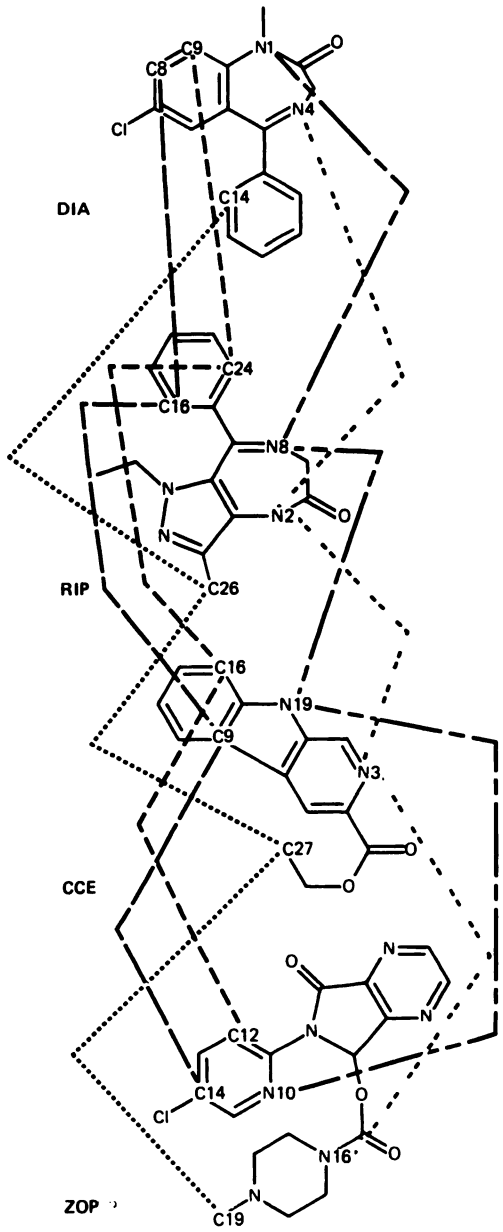

FIG. 2. *Corresponding atoms among the ligands*

The five atoms common to the first 18 ligands are linked by *dashed lines* for the representative molecules DIA, RIP, CCE, and ZOP. The molecular structures lie roughly in the plane of the page in orientations and conformations that maintain the correspondence. MCC follows the pattern of CCE, and the remaining 13 ligands resemble DIA. Hydrogen atoms are not indicated.

that CCE does. More is required than just s10 alone, however. The site is modeled to be a bit flexible (δd = 0.35 Å in this work) in order that slightly differing intramolecular distances will match the corresponding intersite point distances satisfactorily. There follows from this flexibility a small wobble in the fit of even the rigid MCC ring bound at three non-colinear points. Thus atom C30 could evade s10 by wobbling to one side. Placing s14 nearby, as shown in Fig. 3, covers this alternative and forces the repulsion of C30. By similar reasoning, Cl33 of R48 must be responsible for this benzodiazepine analogue's poor binding. Here the wobbling has such a large lever arm that it was necessary to propose s11 as a large spherical disallowed region, blocking 4′ substituents by forming a wall there (Fig. 4). ZOX presents a more subtle problem. In Fig. 1 it appears to have strictly a subset of the atoms present in DIA. The structural difference I chose to exploit is that ZOX is planar, whereas the benzodiazepines are not. If ZOX attempts to bind in a mode analogous to that of DIA, s12 is positioned carefully to force an unfavorable contact with the O4 of ZOX without interfering with the O2 of DIA. THO, THM, TMO, and TME bind quite poorly, although the one saturated and hence puckered ring does not geometrically exclude a binding mode analogous to that of CCE. If, for example, THO has its N22 bound to s2, N4 to s5, and C8 to s8, then the ring puckering swings C15 to quite a different position than that which the C12 of CCE

TABLE 2

Proposed benzodiazepine receptor site

Site point	Binding role[a]	Coordinates, Å			Steric radius
		x	*y*	*z*	
s1	Position 1 nitrogen	−2.358	−0.419	−0.063	1.0
s2	2′ Aromatic C or N	2.941	0.724	0.212	1.0
s3	6′ Aromatic carbon	2.506	0.588	−2.144	1.0
s4	Position 8 carbon	−2.176	3.196	−0.017	1.0
s5	Position 4 nitrogen	0.242	−1.284	−0.957	1.0
s6	Position 2 carbonyl oxygen	−2.501	−2.669	0.125	1.0
s7	Position 2′ substituent	2.628	0.153	1.602	1.0
s8	Bridgehead C (DIA C11)[b]	−0.461	0.978	−0.401	1.0
s9	Position 9 carbon	−2.640	1.962	−0.487	1.0
s10	Steric repulsion of C30, methyl group in MCC[b]	2.565	−1.817	−1.697	1.0
s11	Steric repulsion of Cl33 of R48[b]	16.000	1.000	−1.500	10.0
s12	Steric repulsion of O4 of ZOX[b]	−1.100	−2.250	0.300	0.9
s13	Steric repulsion of THO, etc; atom C15[b] in mode 0 N22 0 0 C8 . . .	3.000	7.000	2.000	3.5
s14	Same as for s10	2.276	−2.364	0.089	1.0
s15	Steric repulsion of atom S20 of PRO[b] in mode 0 C5 C2 0 0 0 Cl40 . . .	4.000	2.000	−20.000	17.8

[a] Position numbering refers to the standard 1,4-benzodiazepine ring-numbering scheme, which is followed in the atom labeling of DIA in Fig. 1.

[b] See atom labeling in Fig. 1.

would occupy. Thus s13 is another large repulsive site point intended to strike C15 of THO and its analogues (Fig. 4). THE binds rather better, presumably owing to favorable interactions with its ethyl group. The poor binding of CHO must be explained away not on steric grounds, but rather that the acidic carboxyl group is less favorable in its interactions than is an ester group. To account for the poor binding of PRO, I observed that the energetically favorable binding modes placed its ring sulfur atom, S20, always beneath the general plane of the site points s1–s9. The very large repulsive sphere s15 is intended to build a floor under the site shown in Fig. 4, where it prevents PRO from binding well without interfering with the binding of all the other ligands. The site point roles are summarized in Table 2, and all proposed binding modes are given in Table 1.

SITE CALCULATION

The standard method of generating site point coordinates is to deduce upper and lower bounds on the intersite point distances from the desired binding modes, and then use the usual distance geometry embedding algorithm to produce coordinates consistent with the bounds (2). An important parameter at this stage is the permissible error in distance range matching, δd, as discussed in the previous section. If δd is large, a wide range of site point coordinates is permissible, and there will be a comparatively large number of geometrically allowed binding modes for each ligand. It is therefore important to keep δd as small as possible in order to hold computation costs to a feasible level. Of course, the costs also increase exponentially with the number of open site points (inclusion of steric repulsion site points is inexpensive).

With these considerations in mind, I began by calculating intersite point distance bounds on the 10 site points s1–s10 described in the previous section, employing the proposed binding modes suggested by the decomposition analysis. The modes were essentially those of Table 1 with the exceptions of those for ZOP, RIP, CCE, and MCC. For example, at this stage it is not necessary to have guessed *all* of the contacts a ligand could make with the site. If a contact is left out of the proposed binding mode, there will simply be possibly less stringent distance bounds produced. The search for the energetically optimal, geometrically allowed binding mode will then discover that an extra contact is possible and energetically favorable. The bound deduction algorithm is that of ref. 2 with the improvement that initially δd is zero, and it is increased only when necessary to make geometrically consistent a proposed contact. By altering the proposed contacts which cause δd to be raised, one can keep its final value relatively low, in this case 0.35 Å, a great improvement over the 1.5 Å needed in the decomposition analysis. Another required modification of the binding modes required in this step is that the mode for MCC, for example, must be the same as that of CCE with the additional contact of the C30 of MCC with site point s10. Only in this way will the distance bounds relating s10 to the rest of the site be produced. The subsequent calculation of energetically optimal binding modes will realize that contacts with the repulsive s10 are intolerable, and alternate modes for MCC will be suggested.

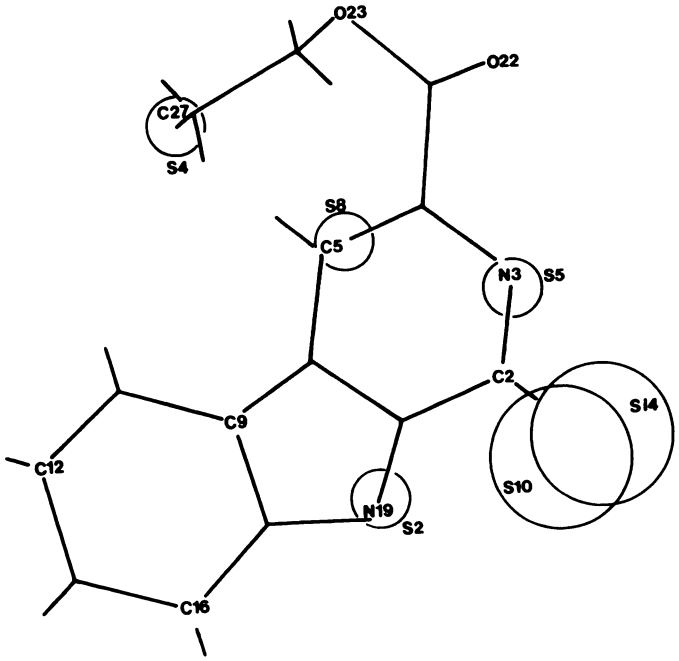

FIG. 3. *Proposed receptor site with CCE bound*

Not all site points are shown for the sake of clarity. Note repulsive site points s10 and s14 that prevent MCC from binding in a similar fashion.

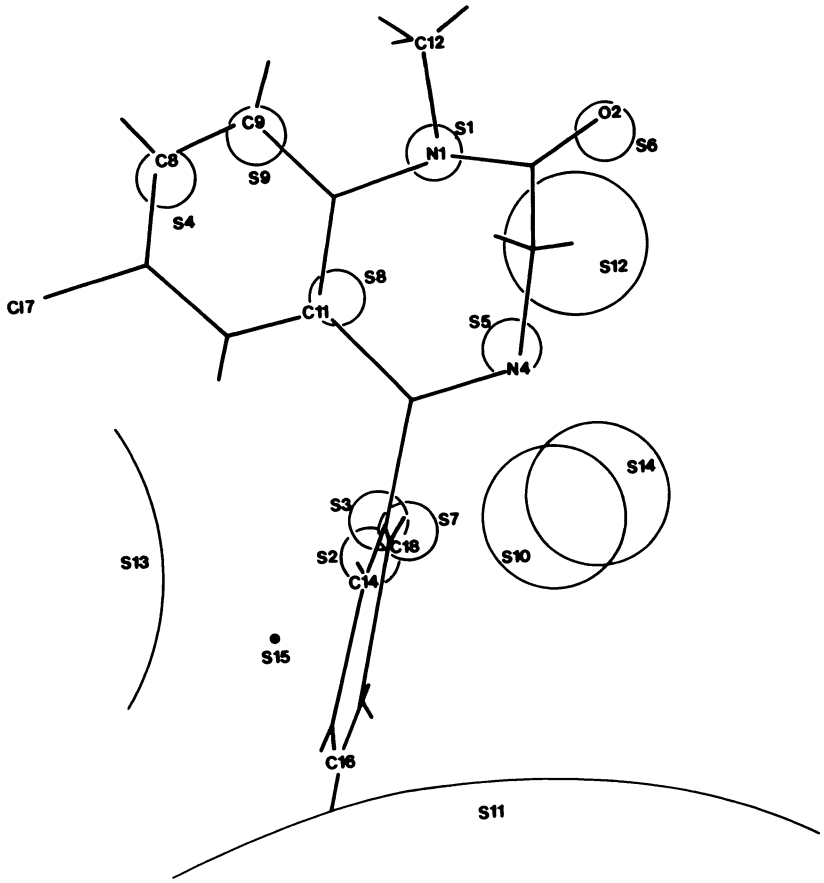

FIG. 4. *Proposed benzodiazepine receptor site with diazepam bound*

Site points s1, s4, s5, s6, s8, s9, and s12 lie approximately in the plane of the page, while s2, s7, and s13 are somewhat above. Point s15 is far below the plane, and s11 and s13 are centered far to the sides, out of view in this illustration. Attractive site points s1–s9 are drawn with radii = 0.35 Å, the tolerance for interatomic distance checks. The repulsive site points s10–s15 have steric repulsion radii given in Table 2. Any atom falling within such a sphere would be sterically disallowed. The large site point s15 presents a broad spherical surface under most of the molecule.

The intersite point distance bounds deduced above allowed variations of only tenths of Angstroms among the coordinate sets produced by the geometry embedding program. The main reason is that there are so many large rigid groups in the ligands, so that the deduced upper bounds nearly equaled their corresponding lower bounds. The coordinates of s11–s15 were found only after the first 10 site points were produced by systematically binding the appropriate low-affinity ligand in a mode to be excluded, calculating the forced coordinates of the offending atom that is supposed to strike the steric repulsion site point, and proposing a center and radius for that site point sphere which contains the coordinates of the offending atom. Figure 4 shows one view of the proposed site with DIA bound.

After the geometry of the site was specified, values for the energy parameters had to be found. the method is the same as before (2), finding a least-squares fit of the calculated binding energies to the observed binding energies subject to the constraint that the desired binding modes also be the energetically optimal ones. The refinement algorithm begins with a least-squares fit of the observed free energies of binding. The resultant energy parameters are used to calculate the energetically optimal binding modes of the ligands. If the optimal mode for a ligand differs from the proposed mode, a linear inequality constraint on the energy parameters is derived to ensure that the unwanted mode has a worse calculated energy than does the proposed mode. The *constrained* least-squares fit is then tried again, and so on. As more and more constraints are added in each cycle of refinement, the root mean square fit of the binding energies gradually worsens. It developed in the course of these calculations that the originally proposed modes for ZOP, RIP, and CCE (based on the decomposition analysis) produced inequalities that eventually led to enormously bad least-squares fits. By changing their proposed modes to the energetically preferred modes that the refinement procedure was trying to eliminate, a much better fit was obtained. Similarly, MCC was allowed to seek its own preferred binding mode that avoided contact with s10, and the other poorly binding ligands were allowed to find their best modes. The result is the list of final proposed binding modes given in Table 1. By using these modes, it is possible to obtain a least-squares fit of the experimental data at the energy parameter values of Table 3, where the proposed binding modes are also the energetically optimal ones. The root mean square deviation between observed and calculated free energies of binding is 1.1 kcal. There are 14 adjustable interaction energies listed in Table 3, but at the solution 8 inequalities are active, and one of the interaction energies is driven to its

TABLE 3

Proposed interaction energy table

In kilocalories per mole. Blank entries indicate that those entries were not used by the desired binding modes and remained fixed at the arbitrary default value of +10 kcal/mole. All interaction energies with site points s10–s15 are +10 kcal/mole.

Ligand point types	Site point								
	s1	s2	s3	s4	s5	s6	s7	s8	s9
H									
C		−0.83	−0.0	−1.48				−2.98	−0.54
O	−1.61					−2.10			
N	−0.05	−2.80			−2.20	−2.00			−2.74
S									
Cl							−1.50		
F							−2.76		
Br									
Carboxylate O^-									

upper limit of zero. Therefore, there are only 14−8−1 = 5 degrees of energy freedom in the fit of 29 observations.

DISCUSSION

The first conclusion of this work is that the distance geometry approach to relating chemical structure to biological activity certainly does not depend on similarity among the ligands. Reasonable results were obtained for a collection of 29 drugs, of which the tightly binding ones belong to 5 different classes of compounds: benzodiazepines, triazolobenzodiazepines, carbolines, pyrazolodiazepine (RIP), and pyrrolopyrazine (ZOP).

The second conclusion is that there is a structural similarity throughout this diverse collection of drugs, namely that shown in Fig. 2. However, there may be other common features involving as many as five nonhydrogen atoms that these calculations did not discover. Furthermore, it is not clear that this one correspondence is the most physically significant one to their mode of pharmacological action.

The third conclusion is that a simple model binding site can be constructed out of only nine attractive points and five energetic degrees of freedom that gives calculated ΔG values of binding in reasonable agreement with the experimental data. The striking qualitative outcome is that a different correspondence in common structural features between DIA and, respectively, ZOP, CCE, and RIP is required for a satisfactory fit of the data. Thus the final proposed binding modes for DIA (and the other benzodiazepines), ZOP, RIP, and CCE given in Table 1 do not agree with the correspondences shown in Fig. 2. This implies that we should be cautious when discovering a structural similarity and immediately proposing that all good ligands bind with corresponding atoms in corresponding positions relative to the site. There may well be many equally appealing correspondences, not all of which can be developed into consistent models for the binding site. In the present case, I suggest that the modes of Table 1 and the corresponding Figs. 3–5 give a truer picture of the way in which these drugs bind to the receptor site. The poorly binding ligands in this study are important for proposing the positions of the repulsive site points, but their modes, shown in Table 1, cannot be considered very significant. Experimentally, all we know

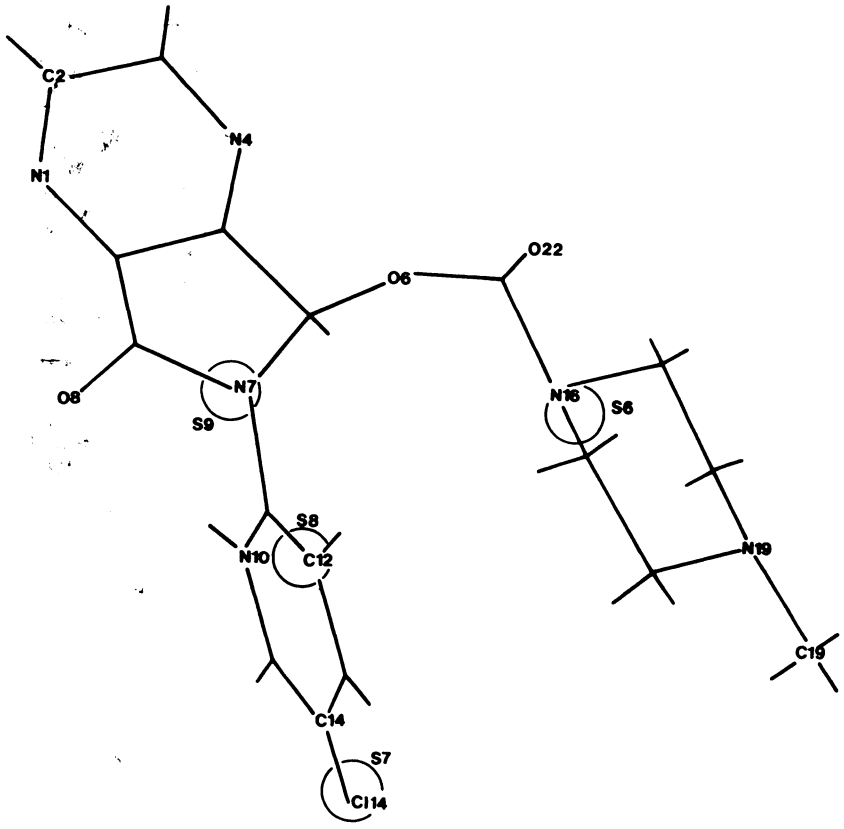

FIG. 5. *Proposed receptor site with ZOP bound*
Not all site points are shown for the sake of clarity. The orientation of the site points is the same as in Figs. 3 and 4.

is that they must not fit well into the site, and the computational result is that there are over 100 binding modes for each ligand, each mode having an interaction energy no more favorable than that quoted in Table 1. In contrast, DIA has some 500 geometrically allowed binding modes, but the desired one is considerably more favorable than the alternatives because only in the desired mode can so many favorable contacts be made simultaneously.

One should be cautious about the proposed model of the site shown in Figs. 3–5. Additional binding data may force substantial changes in it, rather than slight adjustments of energy parameters and site point positions, or rather than adding detail. Although presumably the receptor site really is some sort of pocket with space-filling atoms in most general directions, the large steric repulsion spheres that I have used to account for the data may well be too restrictive, and lack the real receptor's crevices that a cleverly designed ligand could exploit. However, the model is directly testable. ZOP is large and has several rotable bonds, so that there are many rather favorable binding modes available to it. Only the lowest, the "desired," mode is shown in Fig. 5. If this mode is indeed correct, conformationally restricted analogues with appropriate ring closure between N10 and N16, for example, should bind at least as well. The picture further implies that massive substitutions in positions 8 and 9 of the benzodiazepine ring system are sterically allowed, at least if they lie at the angle occupied by the ZOP ring containing N1, as shown in Fig. 5. Conservatively speaking, Fig. 4 points out our ignorance of the receptor site. The region at the top of the illustration may well be completely open to solvent. This data set required no specific interaction with the halogen or nitro substituents on position 7 of the benzodiazepine ring. Although the ring system would seem to lie on the broad surface outlined by s15, there was no necessity to propose any structure above the benzodiazepine ring. Site points s4, s8, and s9 should be interpreted as indications of a hydrophobic slot, rather than very specific atom-atom interactions. This data set leaves us in complete ignorance of the effect of subtitutions as positions 6, 8, and 9. Bulky bridges arching above the plane of the benzo-ring in Fig. 4 from position 6 to position 9 may be sterically permitted and even beneficial. Predicting better drugs (especially of radically different chemical structure) given the site point coordinates in Table 2 and the interaction energies in Table 3 is an exciting problem that I leave for the moment to those blessed with computer graphics equipment.

There are some shortcomings in this work. First, the authors of the original binding data did not give explicit estimates of errors, so it is not known whether my calculated binding energies differ from the experimental ones by more than the experimental error. Small differences in the binding assays used in the three different papers (4, 6, 7) may also contribute to fitting errors. Second, the fit is a least-squares fit, which seeks to minimize the sum of the squares of the deviations, rather than minimizing the greatest absolute deviation, for example. Since much of the experimental data lies in a narrow range about −10.5 kcal, there is a "majority rule" effect that gives a good fit to these data points while neglecting the outliers: CCE and CHM. Nonetheless, the correlation coefficient is reasonably high, +0.89.

In sum, this work has shown how to use the distance geometry approach to deal with a structurally diverse set of ligands and develop a reasonable, quantitative model of the binding site. In particular, the method appears to handle well rather subtle steric effects. I have proposed what may be the significant similarities in these extremely dissimilar anxiolytic drugs. I look forward to including additional high quality, *in vitro* binding data on those compounds that will challenge and improve the model.

ACKNOWLEDGMENTS

I am very grateful to Dr. N. C. Cohen for bringing the benzodiazepines to my attention and for introducing me to the literature on the subject. Drs. Richard Rosenfield and Edgar Meyer kindly aided this work by helping me search the Cambridge Crystallographic Data tapes.

REFERENCES

1. Crippen, G. M. Distance geometry approach to rationalizing binding data. *J. Med. Chem.* **22**:988–997 (1979).
2. Crippen, G. M. Quantitative structure-activity relationships by distance geometry: systematic analysis of dihydrofolate reductase inhibitors. *J. Med. Chem.* **23**:599–606 (1980).
3. Crippen, G. M. Quantitative structure-activity relationships by distance geometry: thyroxine binding site. *J. Med. Chem.* **24**:198–203 (1981).
4. Braestrup, C., and R. F. Squires. Pharmacological characterization of benzodiazepine receptors in the brain. *Eur. J. Pharmacol.* **48**:263–270 (1978).
5. Nielsen, M., and C. Braestrup. Ethyl β-carboline-3-carboxylate shows differential benzodiazepine receptor interaction. *Nature (Lond.)* **286**:606–607 (1980).
6. Blanchard, J. C., A. Boireau, C. Garret, and L. Julou. *In vitro* and *in vivo* inhibition by zopiclone of benzodiazepine binding to rodent brain receptors. *Life Sci.* **24**:2417–2420 (1979).
7. Braestrup, C., M. Nielsen, and C. E. Olsen. Urinary and brain β-carboline-3-carboxylates as potent inhibitors of brain benzodiazepine receptors. *Proc. Natl. Acad. Sci. U. S. A.* **77**:2288–2292 (1980).
8. Sternbach, L. H. The benzodiazepine story. *J. Med. Chem.* **22**:1–7 (1979).
9. Gilli, G., V. Bertolasi, M. Sacerdoti, and P. A. Borea. The crystal and molecular structure of 7-chloro-1,3-dihydro-3-hydroxy-5-phenyl-2H-1,4-benzodiazepin-2-one (oxazepam). *Acta Crystallogr. Sect. B Struct. Crystallogr. Cryst. Chem.* **34**:2826–2829 (1978).
10. Camerman, A., and N. Camerman. Stereochemical basis of anticonvulsant drug action. II. molecular structure of diazepam. *J. Am. Chem. Soc.* **94**:268–272 (1972).
11. Bandoli, G., and D. A. Clemente. Crystal, molecular, and electronic structure of an antianxiety agent: 7-chloro-5-(2-chlorophenyl)-1,3-dihydro-3-hydroxy-1,4-benzodiazepin-2-one. *J. Chem. Soc. Perkin Trans.* **II**:413–418 (1976).
12. Chananont, P., T. A. Hamor, and I. L. Martin. Structure of 7-chloro-2,3-dihydro-1-methyl-5-phenyl-1H-1,4-benzodiazepine hydrochloride (medazepam hydrochloride). *Acta Crystallogr. Sect. B. Struct. Crystallogr. Cryst. Chem.* **36**:898–902 (1980).
13. Herrnstadt, C., D. Mootz, H. Wunderlich, and H. Moehrle. Protonation sites of organic bases with several nitrogen functions: crystal structures of salts of chlordiazepoxide, dihydralazine, and phenformin. *J. Chem. Soc. Perkin Trans.* **II**:735–740 (1979).
14. Groth, P. Crystal structure of some benzoxazoline derivatives. *Acta Chem. Scand.* **27**:945–969 (1973).
15. McDowell, J. J. H. The crystal and molecular structure of chlorpromazine. *Acta Crystallogr. Sect. Struct. Crystallogr. Cryst. Chem. B* **25**:2175–2181 (1969).

Send reprint requests to: Dr. G. M. Crippen, Chemistry Department, Texas A&M University, College Station, Tex. 77843.