by column chromatography (silica gel; benzene) to give 1.52 g (67%) of **9**, mp 91.5–92 °C (from EtOH). Anal. ($C_{21}H_{22}O_2$) C, H.

***N*-Benzyl-3-cyano-*N*-methyl-3-phenylbutylamine (10).** A mixture of α-methyl-α-(β-chloroethyl)phenylacetonitrile (13 g, 67 mmol) and *N*-methylbenzylamine (2 g, 165 mmol) was heated for 2 h at 140 °C, allowed to cool, and then diluted with H_2O. The solution was washed with Et_2O, made basic with 20% aqueous NaOH, and extracted with Et_2O. After the solvent was removed, the residue was distilled in vacuo to give 13.4 g (72%) of **10**, bp 190–195 °C (5 mmHg). Anal. ($C_{19}H_{22}N_2$) C, H, N.

***N*-Benzyl-3-(methoxycarbonyl)-*N*-methyl-3-phenylbutylamine (11).** A mixture of **10** (12 g, 43 mmol), absolute MeOH (20 mL), and concentrated H_2SO_4 (13 g) was heated in a sealed tube for 25 h at 140 °C. After cooling, the mixture was diluted with H_2O, made basic with 20% aqueous NaOH, and extracted with Et_2O. Evaporation of the solvent gave 5.1 g (38%) of **11** as a colorless oil. Anal. ($C_{20}H_{25}NO_2$) C, H, N.

4-[(*N*-Benzyl-*N*-methylamino)ethyl]-4-methylisochroman (12). Compound **11** (3.2 g, 10 mmol) was reduced with $LiAlH_4$ in Et_2O by the usual method to give a colorless paste, which was dissolved in dry dioxane (20 mL). To this solution were added paraformaldehyde (4 g) and concentrated HCl (30 mL). The mixture was allowed to reflux for 10 h with bubbling in dry HCl, allowed to cool, made basic with 20% aqueous NaOH, and extracted with Et_2O. After the solvent was removed, the residue was purified by column chromatography (Al_2O_3; benzene) to give 0.6 g (25%) of **12** as an oil. Anal. ($C_{20}H_{25}NO$) C, H, N.

1′-Benzyl-1′-methyl-1-oxospiro[isochroman-3,4′-piperidinium] Iodide (13). A mixture of **2p** (0.5 g, 1.6 mmol), methyl iodide (2 g), and dry Et_2O (20 mL) was stirred for 24 h at room temperature. The precipitate of **13** which separated out was filtered, washed with Et_2O, and dried in vacuo: mp 168–171 °C; yield 0.7 g (95%). Anal. ($C_{21}H_{24}INO_2$) C, H, N.

Biological Tests. Method I. Male Wister rats weighing 300–350 g were stunned and exsanguinated by cutting the corotid arteries. Ten milliliters of a buffered physiological salt solution [NaCl, 154 mmol; KCl, 2.7 mmol; $CaCl_2$, 0.9 mmol; glucose, 5.6 mmol; and Tris-HCl buffer, 10 mmol (pH 7.4)] were injected into the abdominal cavity, and the abdominal wall was gently massaged for 90 s. The fluid of the abdominal cavity was collected and centrifuged at 100 g for 5 min at 4 °C. The pellet was resuspended in fresh buffered physiological salt solution. The peritoneal cell suspension contained approximately 10% mast cells. The cell suspension (1.8 mL in each test tube) was prewarmed at 37 °C for 5 min, the test compound, dissolved in 0.1 mL of the same buffer solution, was added, and incubation was continued for 15 min. Thereafter, compound 48/80, dissolved in 0.1 mL of buffer solution, was added to obtain a final concentration of 0.25 μg/mL, and incubation was continued for another 15 min. The histamine-releasing process was stopped by chilling the test tube in ice–water; this was followed by centrifugation at 400 g for 5 min at 0 °C. A few drops of 1 N HCl were added to the supernatants and to suspensions of the sediments; they were resuspened in 2 mL of a fresh buffer solution and placed in boiling water for 5 min. Histamine in those samples was extracted and determined by the *o*-phthalaldehyde spectrofluorometric procedure of Shore et al.[7] All measurements were carried out in duplicate. In many cases, the fluorescence of the histamine-OPT fluorophore was not affected by the test compound after repeated extractions. However, in some cases, fluorophore formation was appreciably hampered even after several extractions. In those cases, the recovery of histamine from media containing various amounts of histamine and test compounds was measured beforehand, and the histamine content in the samples was estimated, based on this corrected standard curve. Histamine release was expressed in percent of total histamine content of the cells. The spontaneous release, ranging from 1 to 5%, was subtracted. The percent inhibition of histamine release was calculated according to the formula $[(A - B)/A]100$, where A represents the mean percentage of histamine release induced by compound 48/80 alone and B represents the mean percentage of histamine release elicited by compound 48/80 following pretreatment of the test compound. All data were expressed as mean values ± SEM and each value was calculated from data obtained for at least four rats. The statistical significance of the results was determined by Student's t test ($p < 0.05$). If inhibition of histamine release was less than 10%, the test compound was considered as inactive.

Method II. Mast cells, harvested as described above, were incubated for 15 min at 37 °C in media containing the test compounds in CMC suspension. A few drops of mast cell solution were placed in a 37 °C microbath and exposed for 15 min to compound 48/80 solution (final concentration 0.25 μg/mL). Morphological changes of the mast cells were observed under an inverted phase-contrast microscope (Nikon MD, ×600) and in each case more than 100 mast cells were counted. The percent inhibition of degranulation was calculated, using the formula $100[1 - [(X - Z)/(Y - Z)]]$, where X represents the mean percentage of degranulated cells pretreated with the test compound and then treated with compound 48/80, Y represents the mean percentage of degranulated mast cells exposed to compound 48/80 alone, and Z represents the mean percentage of cells undergoing degranulation in the control medium.

(7) For a method of the fluorometric assay of histamine in tissues, see P. A. Shore, A. Burkhalter, and V. J. Cohn, Jr., *J. Pharmacol. Exp. Ther.*, 28, 127, 182 (1959).

Quantitative Structure–Activity Relationships by Distance Geometry: Thyroxine Binding Site[1]

Gordon M. Crippen

Department of Chemistry, Texas A&M University, College Station, Texas 77843. *Received July 21, 1980*

The binding data for 27 analogues of thyroxine interacting with human prealbumin have been fitted to a simple distance geometry model of the binding site. The novel aspect of the calculation is that a general definition of chirality has been implemented, so that the model site exhibits the observed stereospecificity. The calculated free energies of binding of the ligands fit the experimental data with a root mean square deviation of 0.5 kcal/mol. Three additional analogues not included in the original data set fit equally well. The geometry of the proposed binding site matches the X-ray crystal structure of the tetraiodothyronine–prealbumin complex with a root mean square deviation of 1.0 to 1.6 Å.

This is the third paper in a series on a novel 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[2] discussed the basic simplifying assumptions employed and the format for framing up one's hypothesis about the modes of binding of the ligands. The second article[3] elaborated on additional algorithms de-

(1) This work was supported by National Science Foundation Grant PCM8002602.

(2) Crippen, G. M. *J. Med. Chem.* **1979**, *22*, 988.
(3) Crippen, G. M. *J. Med. Chem.* **1980**, *23*, 599.

0022-2623/81/1824-0198$01.00/0

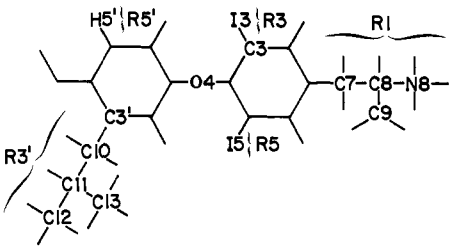

Figure 1. Typical ligand structure (12) showing atom and position labeling used in Tables I, II, and IV. Only labeled atoms are included in the representations used in binding calculations. Unmarked atoms are C, H, 4'-OH, and C9 carboxyl oxygens.

signed to make the construction of a model site more automatic and less laborious. However, in both of these works, the test data involved only optically inactive compounds. In this paper we describe how stereospecific binding has now been included; otherwise, the methodology is that of ref 3.

The test case I chose is the excellent data of Andrea et al.[4] Their precise experiments on the free energy of binding of thyroxine analogues to the same site on pure human prealbumin are exactly the sort of data needed. The structures of the ligands and their binding energies are summarized in Table I and Figure 1. Both the L and D stereoisomers of 3,5,3'-triiodothyronine and tetraiodothyronine are included (ligands **1**, **21**, **2**, and **26** in Table I), so stereospecific binding can be modeled. A further advantage of these data is that the X-ray crystal structure of the ligand **2**–prealbumin complex is known,[5] and preliminary atomic coordinates of the binding site region including the ligand have been kindly made available to me.[6] The binding site of the present paper has been calculated without including evidence from the crystal structure, but rather it has been used only as a final check on our proposed geometry.

Methods

The basic steps are the same as those already described in our previous paper,[3] but with some minor improvements. Calculations were performed on the Texas A&M University data processing center's Amdahl 470V/6 computer. The 27 ligands of Table I were represented initially as each being a collection of points, with one point per atom, including hydrogens. Coordinates for each ligand were derived from the crystal structure of 3,5,3'-triiodo-L-thyronine[7] (ligand 1), deleting or adding groups where necessary, using standard bond lengths and bond angles. The coordinates of all atoms of a ligand represents only one possible conformation, while the subsequent fitting algorithms require a summary of all possible conformations in the form of a matrix of upper and lower bounds on the interatomic distances. This was done by a "brute force" computer program which sets each rotatable bond in the molecule at one of three dihedral angles (0, 120, and 240° from the starting conformation), trying all possible combinations of values for all the bonds. Sterically disallowed conformations were rejected whenever the van der Waals radii of any two noncovalently bonded atoms overlapped, and the interatomic distances in the remaining structures were checked exhaustively to find the global minimum and maximum values. Of course it is preferable to use more than three values of rotation about each single bond, but the computer time required for the larger analogues would be excessive, since there were as many as 11 rotatable bonds (see Figure 1).

The upper and lower bound interatomic distance matrices are a necessary but insufficient summary of the ensemble of conformations a ligand may adopt. An important missing factor for the present work is the chirality, if any. Since a distance matrix is invariant under mirror inversion, the chiral information must be supplied in an ad hoc fashion. Whenever there are four atoms i, j, k, and l whose mutual distances are relatively invariant (I used the critereon: upper bound − lower bound < 0.1 Å), and they are not mutually coplanar, then the four atoms constitute a chiral center. Under this definition, the atoms need not be directly bonded to each other. The coplanarity and the handedness of the chiral quartet of atoms is measured by the signed volume, T, of the tetrahedron they describe: $T = \mathbf{v}_{ij} \cdot (\mathbf{v}_{ik} \times \mathbf{v}_{il})$, where $\mathbf{v}_{ij}$ is the normalized vector from atom i to atom j, etc. If $|T| < 0.15$, I took the quartet to be essentially coplanar and, hence, achiral. Otherwise, the sign of T gives the handedness of the chiral quartet. This algorithm correctly identifies asymmetric carbon atoms, such as C8 in Figure 1, by locating all quartets involving it and its four substituents. However, the definition also includes such quartets as C8, C7, one of the hydrogens attached to C7, and O4 (see Figure 1). This would be entirely correct if the two hydrogens on C7 were distinguishable, but the present program considers each atom to be unique. Aside from this detail, the definition is extremely general and correct, encompassing ordinary asymmetric carbon atoms, spiranes, and even chirality due to steric hindrance.

Representing each atom as a separate point is initially desireable (and feasible), since the upper- and lower-bound distances between all pairs of atoms are determined often by which conformations are sterically disallowed, and these usually involve collisions between hydrogen atoms. However, the next stage of the QSAR analysis calls for comparing the structures of ligands by a lengthy combinatorial matching algorithm that determines the common features (if any) and a short list of the distinctive substituents. Computer time for this depends exponentially on the number of points in each ligand, particularly if, say, the carbon atoms are not distinguished by type but only by relative position in the molecule. We therefore devised an editing program to delete unwanted atoms, along with their corresponding interatomic distances and chiral quartets. The most drastic reduction in points which still preserved the underlying chemical similarity of the analogues and their distinguishing features is outlined in Figure 1. All hydrogens were removed except those at the 3, 5, 3', and 5' positions. Carboxyl, amino, and methyl groups are indicated by single points of distinctive respective types. The 4'-OH was removed, as were all ring carbons except C3 and C3'. Preserving these two ring carbons allows one to detect ring flipping, since

Table I. Binding of Thyroxine Derivatives to Prealbumin

						ΔG, kcal/mol	
no.	R3	R5	R3'	R5'	R1	obsd	calcd
1	I	I	I	H	L-$CH_2CH(NH_2)COOH$	−9.2	−9.6
2	I	I	I	I	L-$CH_2CH(NH_2)COOH$	−10.7	−11.3
3	I	H	I	I	L-$CH_2CH(NH_2)COOH$	−10.0	−9.8
4	I	I	H	H	L-$CH_2CH(NH_2)COOH$	−7.1	−7.0
5	I	H	I	H	L-$CH_2CH(NH_2)COOH$	−7.6	−8.1
6	I	H	H	H	L-$CH_2CH(NH_2)COOH$	−5.2	−5.5
7	I	I	Me	H	L-$CH_2CH(NH_2)COOH$	−7.4	−7.4
8	I	I	*i*-Pr	H	L-$CH_2CH(NH_2)COOH$	−7.8	−8.2
9	I	I	*n*-Pr	H	L-$CH_2CH(NH_2)COOH$	−7.6	−8.2
10	I	I	*s*-Bu	H	L-$CH_2CH(NH_2)COOH$	−9.0	−8.2
11	I	I	*t*-Bu	H	L-$CH_2CH(NH_2)COOH$	−8.0	−8.2
12	I	I	*i*-Bu	H	L-$CH_2CH(NH_2)COOH$	−8.4	−8.2
13	I	I	Bzl	H	L-$CH_2CH(NH_2)COOH$	−7.9	−7.4
14	I	I	Br	H	L-$CH_2CH(NH_2)COOH$	−8.2	−8.2
15	I	I	Cl	H	L-$CH_2CH(NH_2)COOH$	−7.7	−7.7
16	Me	Me	I	H	L-$CH_2CH(NH_2)COOH$	−7.8	−6.9
17	Me	Me	I	I	L-$CH_2CH(NH_2)COOH$	−7.7	−8.6
18	I	I	Br	Br	L-$CH_2CH(NH_2)COOH$	−9.2	−9.2
19	I	I	Me	Me	L-$CH_2CH(NH_2)COOH$	−6.6	−7.4
20	I	I	I	H	CH_2COOH	−9.7	−9.8
21	I	I	I	H	D-$CH_2CH(NH_2)COOH$	−7.0	−7.9
22	I	I	I	I	COOH	−11.1	−10.8
23	I	I	I	I	CH_2COOH	−11.9	−11.5
24	I	I	I	I	$(CH_2)_2COOH$	−11.4	−11.1
25	I	I	I	I	$(CH_2)_3COOH$	−11.3	−11.1
26	I	I	I	I	D-$CH_2CH(NH_2)COOH$	−8.7	−9.6
27	I	I	I	I	$(CH_2)_2NH_2$	−9.7	−9.6

(4) Andrea, T. A.; Cavalieri, R. R.; Goldfine, I. D.; Jorgensen, E. C. *Biochemistry* **1980**, *19*, 55.
(5) Blake, C. C. F.; Oatley, S. J. *Nature (London)* **1977**, *268*, 115.
(6) Blake, C. C. F.; Burridge, J. M.; Oatley, S. J., unpublished results.
(7) Cody, V. *J. Am. Chem. Soc.* **1974**, *96*, 6720.

Table II. Site Point Roles

site point	should bind
1	position 3 substituent, not hydrogen
2	position 5 substituent, not hydrogen
3	position 3′ substituent, not hydrogen
4	position 5′ substituent, not hydrogen
5	position 3′ extra carbon chain
6	position 3′ still further carbon chain
7	C9 carboxyl group in L-thyronine analogues
8	N8 amino group in both L and D analogues
9	C9 carboxyl group in D-thyronine analogues
10	C8 (the asymmetric carbon)
11	C7

they lie unsymmetrically off the respective axes of rotation. The result was that the largest ligands now contained only 14 points, compared to nearly 50 atoms originally. Then the standard "fixed mode" analysis was carried out,[3] first determining the common and individual structural components and then determining an additive contribution of each piece to the total free energy of binding by a least-squares fit to the observed binding energies. Our earlier work had restricted the individual contributions to be less than zero (i.e., favorable), but now we have allowed limited positive values as well by making a suitable change of variable in the input to the quadratic programming algorithm. The reduced number of points improved the computing speed by at least a factor of 30. Leaving out groups that are probably involved in the binding, such as the 4′-OH, is not unrealistic for the purposes of this calculation because they occur in every ligand and, hence, contribute only to the energy associated with the base group. Such structural information is purely redundant and does not affect the calculations.

Given the proposed list of site points and their roles shown in Table II, the next stage is to calculate upper and lower bounds on the intersite distances required by the way the ligands must bind to the site.[3] Here I also had to add the chiral considerations, deducing chiral constraints from the proposed binding modes and the known chiral quartets of the ligands. These were input to the embedding program, along with the distance constraints, as explained in ref 8. The final stage of the embedding is the minimization of a penalty function which contains both terms for deviations from distances and terms for every chiral quartet reflecting their deviations from the desired signed volumes (see equation above defining T). The logic is simply that if all four points of a ligand's chiral quartet are to bind to some given quartet of site points, then these site points must have the same handedness. As a technical aside, I mention that distance constraints among the 11 site points of Table II are insufficient to produce site configurations that match the thyroxine geometry. For example, the C4–O4–C1′ angle between the rings is inadequately specified by distance constraints among site points 1–4 and 11. I overcame this difficulty by introducing extra points to bind O4 and C4′ for embedding purposes only.

The refinement of the energy parameters algorithm was essentially unchanged. Whereas previously all energy parameters not employed by the desired binding modes defaulted to the mildly favorable value −0.1 kcal in order that unforseen modes might be explored during the search for the globally optimal one, I was forced in this work to set them to the arbitrary unfavorable value of +10 kcal. Thus, all the desired binding modes were still allowed, as well as many others, but an even larger number of modes were quickly eliminated since they involved unfavorable energy interactions. The refinement algorithm could still converge to the set of interaction energies which best fit the experimental data while having each ligand's desired binding mode globally energetically optimal, at a considerable saving in computer time. Determining the globally optimal binding mode for given energy parameters was carried out as before by trying all geometrically allowed modes. In this work, the geometric checking now included comparing the chirality of four site points with the handedness of the ligand points bound to them, whenever all four ligand points of a chiral quartet were involved.

Some of the binding data of Andra et al. included the apparent free energy of binding of racemic mixtures. Although I did not use these compounds in deducing the binding site, our final results could be used to calculate a ΔG_{bind} for each of the enantiomers and, thus, an apparent binding constant for the mixture. Let the protein be P and the two ligand enantiomers be D and L. Then at equilibrium, $K_D = [DP]/[D][P]$ and $K_L = [LP]/[L][P]$, while the apparent $K_{obsd} = ([DP] + [LP])/([D] + [L])[P]$. Since the mixture of D and L is exactly racemic, $[D] + [DP] = [L] + [LP]$. It is then straightforward to show that

$$K_{obsd} = \frac{K_D + K_L + 2K_DK_L[P]}{(K_D + K_L)[P] + 2}$$

This formula has the curious property of giving the arithmetic mean of K_D and K_L at low protein concentrations but the harmonic mean for high [P]. In the present case,[4] [P] = 50 nM, so the arithmetic mean predominates. Thus, I calculated the binding energies of both enantiomers, used $\Delta G = -RT \ln K$ to find the individual equilibrium constants, calculated the apparent equilibrium constant according to the equation above, and compared with experiment.

Results

After the ligand atomic coordinates were generated, their interatomic distance bound matrices and chiral quartets were calculated, and all unessential atoms were removed, I began the analysis of their structural variability. Referring to the atom-labeling scheme of Figure 1, I found that the part common to all 27 ligands consisted of atoms C7, C3, O4, and C3′. Other atoms in common had been removed for the sake of computational speed. The differences in structure of the individual ligands could be expressed in terms of 25 chemically and geometrically distinct groups, such as I3, H5′, etc. Due to the 2-Å length of the C–I bond, I3 must be geometrically distinct from H3, since the C–H bond length is 1 Å. In other words, for a site with a small distance tolerance (I used 0.36 Å), there would have to be separate sites for hydrogen and iodine substituents at the same position on the ring. The distance tolerance was enough to allow a single, say, 3-methyl group (C–C bond length 1.5 Å) to bind to a site intended for a 3-iodo (site point 1) because the methyl is represented as a single point at its carbon atom. However, simultaneously binding a 5-methyl to site point 2 (positioned for a 5-iodo group) could not be accomodated. Although ligands **7–13** exhibit a variety of aliphatic and aromatic substituents on the 3′ position, the analysis showed only three geometrically distinct aliphatic groups and the benzyl group. The N8 and C9 groups of the D analogues were flagged as geometrically distinctive, but otherwise the variations on the carboxyl end of the molecule seen in ligands **20**, **22–25**, and **27** corresponded well to atom positions of the usual thyroxine. It was these considerations that led me to propose the 11 site points given in Table II. The least-squares fit of the observed binding energies to a calculated energy, defined as the sum of contributions from each component part of the ligand, yielded an rms deviation of 0.33 kcal/mol and a worst error of 0.8 kcal. (These calculated binding energies are not the ones given in Table I, which were found by the subsequent variable mode analysis.) Removing the worst offenders, ligands **3**, **10**, and **17**, brought the rms deviation down to 0.13 kcal. My conclusion was that assuming the ligands to bind in quite analogous positions would cause an error in the fit of only about 0.3 kcal for this data set.

The roles of the site points are given in Table II. The reasons for our choice come largely from the previous paragraph. Apparently no alternate binding modes are necessary to explain the data. Only two extra site points

(8) Crippen, G. M.; Oppenheimer, N. J.; Connolly, M. L. *Int. J. Pept. Protein Res.*, in press.

Table III. Desired Binding Modes

ligand	site point 1	2	3	4	5	6	7	8	9	10	11
1	I3	I5	I3′				C9	N8		C8	C7
2	I3	I5	I3′	I5′			C9	N8		C8	C7
3	I3		I3′	I5′			C9	N8		C8	C7
4	I3	I5					C9	N8		C8	C7
5	I3		I3′				C9	N8		C8	C7
6	I3						C9	N8		C8	C7
7	I3	I5	C10				C9	N8		C8	C7
8	I3	I5	C11		C10	C12	C9	N8		C8	C7
9	I3	I5	C10		C11	C12	C9	N8		C8	C7
10	I3	I5	C10		C12	C13	C9	N8		C8	C7
11	I3	I5	C11		C10	C12	C9	N8		C8	C7
12	I3	I5	C10		C11	C12	C9	N8		C8	C7
13	I3	I5	C10				C9	N8		C8	C7
14	I3	I5	Br3′				C9	N8		C8	C7
15	I3	I5	C13′				C9	N8		C8	C7
16	C10		I3′				C9	N8		C8	C7
17	C10		I3′	I5′			C9	N8		C8	C7
18	I3	I5	Br3′	Br5′			C9	N8		C8	C7
19	I3	I5	C10				C9	N8		C8	C7
20	I3	I5	I3′				O8A			C8	C7
21	I3	I5	I3′					N8	C9	C8	C7
22	I3	I5	I3′	I5′						O7A	C7
23	I3	I5	I3′	I5′			O8A			C8	C7
24	I3	I5	I3′	I5′			C9			C8	C7
25	I3	I5	I3′	I5′			C9			C8	C7
26	I3	I5	I3′	I5′				N8	C9	C8	C7
27	I3	I5	I3′	I5′				N8		C8	C7

around the 3′ position are required for all the aliphatic substituents there, and a site point for I3 cannot be expected to accomodate the H3 of another ligand for geometric reasons. It is not clear how the asymmetric carbon should be treated, so I arbitrarily decided that C7, C8, and N8 should always occupy the same respective site points and that there would therefore be one site point for C9 in the L isomers and another for the D isomers. Certainly some kind of site had to be present for C7, C8, C9, and N8 in order to exhibit chirality. Since more complicated sites require longer computation time, the policy was to try to make do with the smallest number of site points. The resultant detailed desired binding mode of each ligand is shown in Table III (ligand point labeling as in Figure 1). The distance tolerance was set to 0.36 Å by trial and error as the smallest value that would permit these desired binding modes.

My standard refinement of interaction energies, subject to the constraints that the desired binding modes also be the energetically preferred ones, converged to a reasonably small rms deviation of 0.50 kcal/mol between calculated and observed binding energies. The calculated binding free energies, given in Table I, have a correlation coefficient of 0.95 with the observed values. The worst errors (0.9 kcal in absolute value) were for the D stereoisomers, ligands **21** and **26**. Apparently the simplified treatment of the chiral center that assumed only that there are alternate binding sites for the C9 carboxyl group is insufficient to explain the observed binding energies. We are confident that a more complicated site, for example, one with alternate site points for both N8 and C9, would allow a better fit to the data. At the solution reached by the quadratic programming algorithm that determines the energy parameters, the energy difference between D and L forms is limited in a complicated fashion by two inequalities to be less than 1.7 kcal (compare ligands **2** and **26**, for instance). These inequalities are required in order to maintain the desired binding mode of all the ligands. The other major errors are for either 3,5-dimethyl or 3′,5′-dimethyl substituted thyronines. Both methyls cannot bind at once in sites placed for iodine because of the differences in bond length. Perhaps adding extra binding sites for 3-, 5-, 3′-, or 5′-methyl groups would improve the fit, but this would have its penalty in increased computing costs. One might ask why the much simpler Free–Wilson ("fixed mode") analysis given above yields a lower rms deviation (0.3 kcal) than this more complicated variable binding mode approach does (0.5 kcal). The difference lies in the number of adjustable energy parameters: 25 in the former case, but only 11 in the latter. Although there are 18 nondefault interaction energies given in Table V, there were 6 active inequalities at the refinement solution, and one of the variables was driven to zero, its limit. In any event, my fit to the experimental data is somewhat worse than the experimental errors given by Andrea et al.,[4] which were typically ±0.1 kcal or as much as ±0.3 kcal for ligands **11** and **13**.

I have explained how the site geometry is determined by deducing distance constraints from the desired binding modes.[3] If the modes are incorrectly chosen, there may be no satisfactory site geometry. In more fortunate circumstances, there may be a unique site or, more generally, there will be some geometric degrees of freedom left undetermined. For the thyroxine binding site, the distance between site points 1 and 3 could range at least as far as 5.7 to 8.3 Å. In terms of the conformation of thyroxine bound to such sites, this corresponds to substantially different relative orientations of the two aromatic rings. Of course the crystal structure of triiodothyronine[7] lies within the range. From the binding data alone it is impossible to decide which orientation to choose, since the desired binding modes are equally compatible with any choice, and then the energy refinement would presumably proceed equally well to the same result. Therefore, out of a random sample of five generated site geometries, we chose the coordinate set (see Table IV) having a 1–3 distance of 6.1 Å because it was the best fit to the protein–ligand crystal structure[6] value of 6.2 Å. This gives an rms deviation of only 1.04 Å between the distance matrices of the site and the 3,5,3′,5′-tetraiodo-L-thyronine molecule

Table IV. Final Proposed Prealbumin Binding Site Coordinates

point no.	coordinates, Å		
	x	y	z
1	−0.757	−1.517	−3.027
2	0.254	−3.513	2.638
3	3.096	2.103	−0.072
4	6.271	−2.857	0.256
5	3.676	2.297	−1.778
6	2.777	3.865	−1.581
7	−3.285	0.889	2.195
8	−3.653	1.863	0.258
9	−5.069	0.466	−0.565
10	−4.390	0.557	1.151
11	−4.250	−0.820	0.959

bound to prealbumin, considering only the eight site points which directly correspond to atoms in the bound molecule (site points 1–4, 7, 8, 10, and 11). The agreement is shown graphically in Figure 2. In all fairness, the worst fit of the five sites I generated, having 1–3 distance of 8.3 Å, gave an rms distance matrix deviation to the crystal structure of 1.63 Å. Incidentally, the orientation of C8, N8, and C9 relative to the C1–C6 ring is also undetermined by the binding data, since all ligands are equally flexible in that region. This, however, does not give rise to large variability in intersite point distances, due to the small size of the R1 group.

It is clear that the distance geometry approach to QSAR is able to provide a reasonable fit to the observed binding energies and is capable of agreeing with the structural evidence insofar as the binding data can imply geometric detail. What is lacking is a test of its predictive power. The 27 ligands in Table I were not all the data given by Andrea et al. The unsubstituted L-thyronine, where R3 = R5 = R3′ = R5′ = H, has an experimentally determined binding energy of "greater than −4.1 kcal/mol", and I calculate −4.0 kcal. For the racemic mixture of 3′,5′-diiodothyronine, we calculate −8.3 kcal for the L form and −6.6 for the D, yielding an apparent calculated ΔG = −7.9 kcal, compared to the experimentally determined −8.6 kcal. For the racemic mixture of 3′-iodothyronine, I similarly calculated −6.6 for the L and −4.9 for the D, corresponding to an apparent −6.2 kcal, neatly matching the experimental value of −6.2 kcal. In other words, my predictive errors for these ligands of the same general class are no worse than our fit to the original data set. The predicted binding mode is the usual one, analogous to that of ligand 1.

The proposed site seems physically reasonable. The interaction energies of Table V are all believably small, the largest being −2.6 kcal/mol for iodine binding to site point 3. Otherwise, site points 1–4 seem energetically similar for the remaining determined entries. Site point 3 shows

Figure 2. Comparison of the X-ray crystal structure of the 3,5,3′,5′-tetraiodo-L-thyronine and prealbumin complex (light lines) to the calculated site in Table V (heavy lines). The lines between site points are shown only for the sake of clarity. This stereoview shows the X-ray crystal coordinates of nonhydrogen atoms of the residues of prealbumin that form the active site. Some side chains pointing away from the ligand have been deleted for clarity. The ligand is located in the center with the 4′-OH at the top and the asymmetric carbon at the bottom. Approximately superimposed on the ligand is the calculated site, made visible by heavy lines drawn between the points. The right arm links site points 11, 1, and 4; the left arm links site points 11, 2, 3, 5, and 6.

a uniform trend of interaction with the halogens, becoming more favorable with increasing size. A much larger binding energy is attributed to C7 attaching to site point 11 than for other aliphatic carbons binding elsewhere. However, one should note that binding C7 largely commits the entire adjacent ring to bind in the real site, although I have provided no special site point in our model for ring carbons. From Figure 2 it is clear that the geometry of my site is reasonable. As I have already explained, the points match the conformation of tetraiodothyronine by design. In addition, there appears to be sufficient room for all the site points while simultaneously fitting the X-ray coordinates of the ligand. Any commentary about the appropriateness of various side chains near the ligand in the site applies equally well to the site points I have calculated, since the geometries match so closely. It is unwarrented to discuss in detail the differences in positioning of the chiral center in the crystal structure vs. my corresponding site points because of possible errors in the crystal structure on such fine details and because of acknowledged shortcomings in the accuracy of our site in that region.

Conclusions

Distance geometry QSAR has been shown to yield correct information about binding site geometry and interaction energies in terms of fitting the observed binding energies, matching the experimentally determined structure of the binding site, and predicting the binding of ligands outside the original data set. Stereospecific binding can now be correctly treated in the calculations. Note, however, that the method does not introduce more detail than is warrented. For example, I make no suggestions about the effect of, say, fluoro derivatives of thyronine, since there were no fluorine atoms in the entire data set.

Table V. Final Proposed Interaction Energy Table[a]

ligand point types	site points										
	1	2	3	4	5	6	7	8	9	10	11
H (aromatic)											
C (aliphatic)	−0.3		−0.4		−0.4	−0.4				−0.5	−1.5
O							−2.2			−2.0	
N (amino)								−0.2			
S (unused)											
F (unused)											
Cl			−0.7								
Br			−1.2	−1.0							
I	−1.5	−1.5	−2.6	−1.7							
C (aromatic)											
C (carboxyl)							−1.7		−0.0		

[a] In kcal/mol. 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/mol.

Since all the ligands were as conformationally flexible as thyroxine, I can not claim to have determined the relative orientation of the two aromatic rings for the ligand as it lies in the site. In general, if the data set is limited in range of binding energies, conformational differences, and chemical structural differences, then necessarily the outcome of this method must be a limited picture of the site. As more information is added, even in the relatively obscure form of structure and binding energy, our image of the binding site must come into focus.

Acknowledgment. Preparation of Figure 2 was made possible by the generous help of Drs. Peter Kollman, Edgar Meyer, Jeff Blainey, and especially Stan Swanson.

Notes

Synthetic Enterobactin Analogues.[1] Carboxamido-2,3-dihydroxyterephthalate Conjugates of Spermine and Spermidine

Frederick L. Weitl,[2a] Kenneth N. Raymond,*[2b] and Patricia W. Durbin[2c]

Divisions of Materials and Molecular Research and Biology and Medicine, Lawrence Berkeley Laboratory, and Department of Chemistry, University of California, Berkeley, California 94720. *Received April 14, 1980*

Two examples of a new class of synthetic polycatecholate ligands, the carboxamido-2,3-dihydroxyterephthalate conjugates of spermine (8) and of spermidine (10), have been synthesized via the generally useful synthon methyl-2,3-dimethoxyterephthaloyl chloride (6). Initial biological evaluation reveals tetrameric terephthalate (8) to be an extremely effective agent for sequestering and removing plutonium from mice; a single 25-μmol/kg (ip) dose of 8 removed 73% of the plutonium citrate previously injected (iv, 1 h earlier). Under the same conditions, trimeric terephthalate (10) excreted only 49% of injected plutonium. In vitro kinetic experiments have shown that 10 rapidly and quantitatively removed Fe from human transferrin. These results are discussed in relation to the design of metal-ion specific sequestering agents.

We have previously described two related research programs for the design and synthesis of specific sequestering agents for iron(III)[3,4] and actinide(IV) metal ions.[4-8] In the case of iron, since the body lacks any mechanism for removing excess amounts of this essential element, it can be an acute or chronic poison. A major program is underway for the development of iron chelating agents to be used in treating Cooley's anemia, a genetic disease which results in chronic iron overload.[9] Our ferric-ion chelating agents are modeled after the siderophores, a class of low-molecular-weight iron sequestering and transport agents that are produced by microbes. The most powerful natural iron chelator known is enterobactin.[10] Since this siderophore incorporates catechol chelating agents (in the form of 2,3-dihydroxybenzoyl groups, DHB), our initial approach has been the incorporation of several substituted DHB groups into multidentate chelate molecules.

Of the radioactive isotopes produced as byproducts of the nuclear fuel cycle, the major long-term radiation hazard is posed by the transuranium actinides. Of the actinides, plutonium is a particularly dangerous biological hazard because of the chemical and biological similarities of Pu(IV) and Fe(III).[11-14] Incorporated plutonium is bound by transferrin, the mammalian iron-transport protein, at the same site that normally binds Fe(III) and is then concentrated in iron storage sites, where most of it remains indefinitely. In order to prepare specific sequestering agents for Pu(IV) and other actinide ions, we have explicitly recognized this similarity of Pu(IV) and Fe(III) in using as chemical models the microbial chelating agents which are so specific for Fe(III).

As direct analogues of the siderophores such as enterobactin[10] and a threonine conjugate of spermidine isolated by Tait,[15] we have prepared tetrameric[5] and trimeric[16] 2,3-dihydroxybenzoyl conjugates incorporating certain linear, cyclic, and platform amines. Direct sulfonation of these compounds produced the 5-sulfonato-2,3-di-

(1) This is paper number 5 in the series "Ferric Ion Sequestering Agents", and also number 5 in the series "Specific Sequestering Agents for the Actinides". For previous papers in these series see ref 3 and 6, respectively.
(2) (a) Materials and Molecular Research Division, Lawrence Berkeley Laboratory. (b) Address correspondence to this author at the Department of Chemistry, University of California, Berkeley. (c) Biology and Medicine Division, Lawrence Berkeley Laboratory.
(3) Weitl, F. L.; Harris, W. R.; Raymond, K. N. *J. Med. Chem.* **1979**, *22*, 1281–1283.
(4) Raymond, K. N.; Harris, W. R.; Carrano, C. J.; Weitl, F. L. *Adv. Chem. Ser.*, in press.
(5) Weitl, F. L.; Raymond, K. N.; Smith, W. L.; Howard, T. R. *J. Am. Chem. Soc.* **1978**, *100*, 1170–1172.
(6) Durbin, P. W.; Jones, E. S.; Raymond, K. N.; Weitl, F. L. *Radiat. Res.* **1980**, *81*, 170–187.
(7) Raymond, K. N.; Smith, W. L.; Weitl, F. L.; Durbin, P. W.; Jones, E. S.; Abu-Dari, K.; Sofen, S. R.; Cooper, S. R. *ACS Symp. Ser.*, in press.
(8) Smith, W. L.; Raymond, K. N. *Struct. Bonding*, in press.
(9) Anderson, W. F.; Heller, M. C., Eds. *DHEW Publ. (NIH) (U.S.)* **1975**, *NIH 77-994*.
(10) Harris, W. R.; Carrano, C. J.; Cooper, S. R.; Sofen, S. R.; Avdeef, A. E.; McArdle, J. V.; Raymond, K. N. *J. Am. Chem. Soc.* **1979**, *101*, 6097–6104.
(11) Durbin, P. W. *Health Phys.* **1975**, *29*, 495–510.
(12) Stoves, B. J.; Atherton, D. R.; Buster, D. S. *Health Phys.* **1971**, *20*, 369.
(13) Catsch, A. "Radioactive Metal Mobilization in Medicine"; Charles C. Thomas: Springfield, Ill., 1964.
(14) Foreman, H.; Moss, W.; Langheim, W. *Health Phys.* **1960**, *2*, 326.
(15) Tait, G. H. *Biochem. J.* **1975**, *146*, 191–204.
(16) Weitl, F. L.; Raymond, K. N. *J. Am. Chem. Soc.* **1979**, *101*, 2728–2731.
(17) Weitl, F. L.; Raymond, K. N. *J. Am. Chem. Soc.*, in press.
(18) Carrano, C. J.; Raymond, K. N. *J. Am. Chem. Soc.* **1979**, *101*, 5401–5404.

0022-2623/81/1824-0203$01.00/0