Molecular Mechanics Simulation of Protein-Ligand Interactions: Binding of Thyroid Hormone Analogues to Prealbumin

Jeffrey M. Blaney,[†] Paul K. Weiner,[†] Andrew Dearing,[†] Peter A. Kollman,*[†] Eugene C. Jorgensen,[†§] Stuart J. Oatley,[‡∥] Jane M. Burridge,[‡] and Colin C. F. Blake[‡]

Contribution from the Department of Pharmaceutical Chemistry, School of Pharmacy, University of California, San Francisco, California 94143, and Laboratory of Molecular Biophysics, University of Oxford, South Parks Road, Oxford OX1 3PS, England.
Received September 2, 1981

Abstract: We have simulated the interaction of L-thyroxine (**1**), D-thyroxine (**2**), and their deamino (**3**) and decarboxy (**4**) analogues with the human plasma protein prealbumin by using molecular mechanics calculations. Starting geometries were taken from the high-resolution X-ray structure of prealbumin and difference electron density maps of the prealbumin-thyroxine complex. We model the interactions by using the atoms of the thyroxine analogue and approximately 250 atoms within the binding site of prealbumin, minimizing the total energy with respect to all geometric degrees of freedom. Using the molecular mechanics calculated interaction energies and a simple empirical method to estimate the solvation energy differences of **1-4**, we qualitatively reproduce the experimentally observed relative free energies of association of these analogues to prealbumin and offer a structural and energetic model to account for the different binding affinities of analogues **1-4** to the protein.

Introduction

One of the most challenging goals in theoretical biochemistry is the development of quantitative models for the forces governing molecular association. If the association is covalent (e.g., an enzyme-substrate interaction), the development of a model for this process is made very difficult by our lack of understanding of the detailed energetics of all but the simplest chemical reactions. However, important progress in the development of such models has been made by Warshel et al.[1] Clementi,[2] and DeTar.[3] In principle, the development of models for noncovalent association is easier since our ability to model these noncovalent forces quantitatively is much greater than our ability to model covalent interactions. It is generally accepted that such forces are hydrophobic, van der Waals, hydrogen bonding, and electrostatic (ion-pairing) interactions.[4] However, the related problems of the immense number of degrees of freedom in a protein-ligand complex and the lack of an adequate model for dealing with solvent make the modeling of noncovalent association extremely difficult. Important in the development of such models has been the work of Platzer et al.,[5] Pincus et al.,[6] Levitt,[7] and Case and Karplus[8] at the "all atom" level and Wodak et al.[9] at a simpler level.

Our interest in the past few years has been focused on the interaction between prealbumin and thyroxine (T_4) analogues due to the possible relevance to the biologically important thyroid hormone-nuclear receptor interaction[10-12] and because these complexes are well suited to developing models for noncovalent protein-ligand association. Human plasma thyroxine-binding prealbumin (TBPA) was the first fully characterized hormone-binding protein. It is a nonglycosidic linear protein consisting of four identical subunits, each containing 127 amino acids. The subunits are associated in an ellipsoidal shape, forming a cylindrical channel containing two identical hormone binding sites. Each binding site has a 2-fold axis of symmetry and contains two pairs of symmetry-equivalent lysine and glutamate residues (Lys-15A, Lys-15C, and Glu-54A, Glu-54C). The residues in the binding site form close contacts with both aromatic rings of T_4, with the carboxyl and ammonium groups of the thyroid hormone associated by ion pairing with the Lys-15 and the Glu-54 residues in the funnel-shaped mouth of the site (Figure 1).[13,14]

It is crucial to have a high-resolution structure of the protein, at least qualitative structural data on the protein-ligand complex, and some experimental free energies of ligand association before developing a theoretical model of protein-ligand association. All of these prerequisites are met by the TBPA-T_4 system, which has the added advantage of a very deep and well-defined binding cleft in the protein and limited conformational flexibility in the ligand. We have thus developed a structural and energetic model for the association of L-T_4 (**1**), D-T_4 (**2**), deamino-T_4 (**3**), and decarboxy-T_4 (**4**) (Figure 2) with TBPA using molecular mechanics. Using the molecular mechanics calculated interaction energies and a simple empirical model to estimate the solvation energy differences of **1-4**, we are able to qualitatively reproduce the experimentally observed free energies of association of these analogues to TBPA ($\Delta G_3 < \Delta G_1 < \Delta G_2 < \Delta G_4$)[15] and offer a structural and energetic model to account for the different binding affinities of these analogues.

Methods

To estimate the binding energies of the T_4 analogue-TBPA complexes, consider the following thermodynamic cycle:

$$\begin{array}{ccc} P_g + A_g & \xrightarrow{\Delta G_g} & PA_g \\ \downarrow \Delta G_{SP} & \downarrow \Delta G_{SA} & \downarrow \Delta G_{SC} \\ P_{aq} + A_{aq} & \xrightarrow{\Delta G_{aq}} & PA_{aq} \end{array}$$

[†] University of California.
[‡] University of Oxford.
[§] Deceased.
[∥] Present address: Department of Pharmaceutical Chemistry, School of Pharmacy, University of California, San Francisco, CA 94143.

(1) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227. Warshel, A. *Proc. Natl. Acad. Sci., U.S.A.* **1978**, *75*, 5250.
(2) Clementi, E. *Int. J. Quant. Chem.* **1980**, *17*, 651.
(3) DeTar, D. *J. Am. Chem. Soc.* **1981**, *103*, 107; *Biochemistry* **1981**, *20*, 1730.
(4) Fersht, A. "Enzyme Structure and Mechanism"; W. H. Freeman: San Francisco, 1977.
(5) Platzer, K. E. B.; Momany, F. A.; Scheraga, H. A. *Int. J. Peptide Protein Res.* **1972**, *4*, 187.
(6) Pincus, H. R.; Burgess, A. W.; Scheraga, H. A. *Biopolymers* **1976**, *15*, 2485.
(7) Levitt, M. *J. Mol. Biol.* **1974**, *2*, 393.
(8) Case, D.; Karplus, M. *J. Mol. Biol.* **1978**, *123*, 697.
(9) Wodak, S.; Janin, J. *J. Mol. Biol.* **1978**, *124*, 323.
(10) Eberhardt, N. L.; Ring, J. C.; Johnson, L. K.; Latham, R. K.; Apriletti, J. W.; Kitsis, R. N.; Baxter, J. D. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 5005.
(11) Eberhardt, N. L.; Ring, J. C.; Latham, K. R.; Baxter, J. D. *J. Biol. Chem.* **1979**, *254*, 8534.
(12) Jorgensen, E. C.; Andrea, T. A.; Bolger, M. B.; Kuo, G. T. J.; Baxter, J. D.; Eberhardt, N. L. *Thyroid Res. 8 [Eight], Proc. Int. Thyroid Congr.* **1980**, abstract.
(13) Blake, C. C. F.; Oatley, S. J. *Nature (London)* **1977**, *268*, 115.
(14) Blake, C. C. F.; Geisow, M. J.; Oatley, S. J.; Rērat, B.; Rērat, C. *J. Mol. Biol.* **1978**, *121*, 339.
(15) Andrea, T. A.; Cavalieri, R. R.; Goldfine, I. D.; Jorgensen, E. C. *Biochemistry* **1980**, *19*, 55.

0002-7863/82/1504-6424$01.25/0

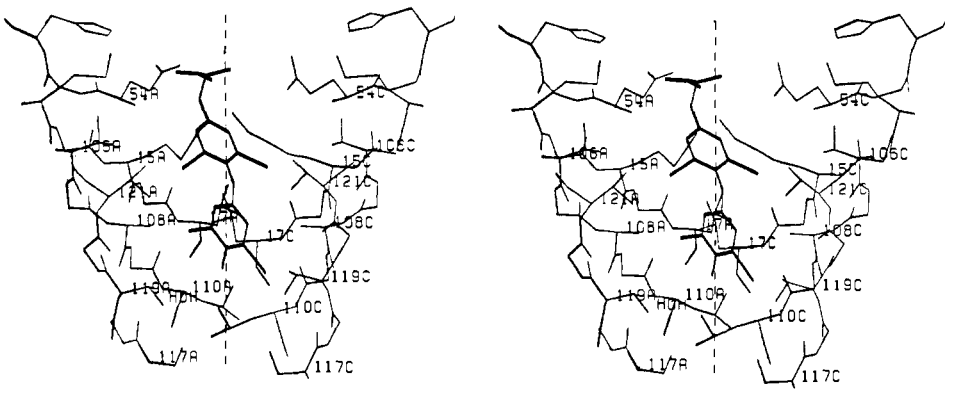


Figure 1. L-T_4 in the prealbumin hormone-binding site (original unmodified coordinates). The dashed line indicates the 2-fold symmetry axis. Labeled residues include Lys-15, Leu-17, Glu-54, Thr-106, Ala-108, Leu-110, Ser-117, Thr-119, Val-121, and the bound water (HOH).

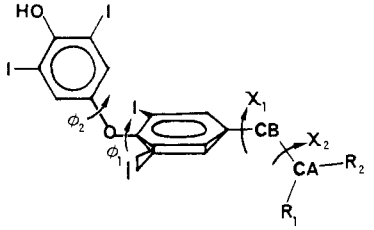


Figure 2. Thyroxine analogues: (**1**) L-T_4, $R_1 = COO^-$, $R_2 = NH_3^+$; (**2**) D-T_4, $R_1 = NH_3^+$, $R_2 = COO^-$; (**3**) deamino-T_4, $R_1 = H$, $R_2 = COO$; (**4**) decarboxy-T_4, $R_1 = NH_3^+$ $R_2 = H$.

The free energy of binding of T_4 analogues to TBPA in aqueous solution (ΔG_{aq}) is related to the free energy of solvation of TBPA (ΔG_{SP}), the free energy of solvation of the analogue (ΔG_{SA}), the free energy of solvation of the protein–ligand complex (ΔG_{SC}), and the free energy of the gas-phase binding of the analogue to TBPA (ΔG_g) by

$$\Delta G_{aq} = \Delta G_g + \Delta G_{SC} - \Delta G_{SP} - \Delta G_{SA} \quad (1)$$

We are interested in $\Delta(\Delta G_{aq})$, the relative binding free energies of the T_4 analogues compared to L-T_4:

$$\Delta(\Delta G_{aq}) = \Delta(\Delta G_g) + \Delta(\Delta G_{SC}) - \Delta(\Delta G_{SP}) - \Delta(\Delta G_{SA}) \quad (2)$$

We make the assumption that the relative free energies of *gas-phase* association $\Delta(\Delta G_g)$ can be approximated by the relative internal energies $\Delta(\Delta G_g) \approx \Delta(\Delta E_g)$ since the various rotational, vibrational and translational entropic and enthalpic contributions to binding should be approximately equal for closely related analogues. Thus

$$\Delta(\Delta G_{aq}) \approx \Delta(\Delta E_g) + \Delta(\Delta G_{SC}) - \Delta(\Delta G_{SP}) - \Delta(\Delta G_{SA}) \quad (3)$$

The values for $\Delta(\Delta E_g)$ are derived from molecular mechanics calculations using the AMBER software package[16,17] on a PDP-11/70 in which the energy of the system is represented in terms of bond-stretching and -bending, torsional, and nonbonded (van der Waals and electrostatic) energies:

$$E_{total} = \sum_{bonds} K_r(r - r_{eq})^2 + \sum_{angles} K_\theta(\theta - \theta_{eq})^2 + \sum_{dihedrals} \frac{K}{2}\{1 + \cos(n\varphi - \gamma)\} + \sum_{i<j} \frac{B_{ij}}{R_{ij}^{12}} - \frac{A_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{\epsilon R_{ij}} \quad (4)$$

All atoms were allowed to move during energy refinement, by using analytical gradients with conjugate gradient minimization until the root-mean-square (rms) energy gradient was less than 0.1 kcal/Å. The force field we employ in eq 4 is similar to that of Gelin and Karplus,[18] differing mainly in our explicit inclusion of potentially hydrogen-bonding hydrogens (bonded to N, O, and S), and is summarized in Table III and Figures 7 and 8. As in ref 18, CH, CH_2, and CH_3 groups are treated as united atoms, except for the aromatic C–H atoms on the T_4 analogues. The partial atomic charges used in evaluating the electrostatic energy come from our previous study of electrostatic potentials of proteins.[19] An "improper" torsional angle (also used to maintain planarity in peptide bonds) was added to the Glu-54 and T_4 carboxyl groups. Electrostatic energies are evaluated by using a dielectric constant of $\epsilon = R_{ij}$ (interpreted as the numerical, dimensionless value of the distance when nonbonded atoms i and j are separated by R_{ij} Å). Warshel was shown that the effective dielectric constant for short-range ionic interactions in water is smaller than the bulk dielectric and increases (roughly linearly) with the ionic separation.[20] The preliminary study of Rees on the internal effective dielectric constant of cytochrome c also suggests that the dielectric constant within proteins increases with charge separation.[21] This dielectric model has also been reasonably successful in molecular dynamics calculations of proteins.[22] By weighting short-range electrostatic interactions more strongly than long range, use of a distance-dependent dielectric constant indirectly allows for polarization and charge-transfer contributions to the contact ion-pair interaction (as opposed to the solvent-separated ion pair). All structures were also refined with a dielectric constant $\epsilon = 2.0$ to compare with the models derived with a distance-dependent dielectric constant. A dielectric constant of 2.0 was chosen because it led to ionic interactions of roughly the same strength as found with $\epsilon = R_{ij}$.

It is not possible to represent solvent explicitly in our models due to computational limitations, so we seek a simple model that will allow us to estimate the *differential* solvation contributions to the free energy of binding for the various analogs, $\Delta(\Delta G_{SC}) - \Delta(\Delta G_{SP}) - \Delta(\Delta G_{SA})$. The simplest approach is to assume that the solvation energies for TBPA and the TBPA–ligand complex are similar for each analogue, thus both $\Delta(\Delta G_{SP})$ and $\Delta(\Delta G_{SC})$ are zero. This is rigorously correct for $\Delta(\Delta G_{SP})$, since the solvation of TBPA itself is independent of the analogue. The validity of the approximation that $\Delta(\Delta G_{SC})$ is zero depends on each analogue being equally "buried" in the TBPA binding site, making the solvation energy of the TBPA–ligand complex independent of the analogue. Therefore

$$\Delta(\Delta G_{aq}) \approx \Delta(\Delta E_g) - \Delta(\Delta G_{SA}) \quad (5)$$

Since the analogues considered here differ only in the sidechain portion of the molecule, we estimate $\Delta(\Delta G_{SA})$ from the free energies of transfer from water to the gas phase of glycine (a model for L- and D-T_4), acetic acid (for deamino-T_4), and methylamine (for decarboxy-T_4). Wolfenden has reported the water/vapor distribution coefficients for the neutral forms of acetic acid[23] (8.1×10^4) and methylamine[24] (3.4×10^3). The water/vapor distribution coefficient for the uncharged form of glycine has been

(16) Weiner, P.; Kollman, P. *J. Comp. Chem.* **1981**, *2*, 287.
(17) Kollman, P.; Weiner, P.; Dearing, A. *Ann. N.Y. Acad. Sci.* **1981**, *367*, 250.
(18) Gelin, B.; Karplus, M. *Biochemistry* **1979**, *18*, 1256.
(19) Hayes, D. M.; Kollman, P. *J. Am. Chem. Soc.* **1976**, *98*, 3335.
(20) Warshel, A. *J. Phys. Chem.* **1979**, *83*, 1640.
(21) Rees, D. *J. Mol. Biol.* **1980**, *141*, 323.
(22) McCammon, J. A.; Wolynes, P. G.; Karplus, M. *Biochemistry* **1979**, *18*, 927.
(23) Wolfenden, R. *J. Am. Chem. Soc.* **1976**, *98*, 1987.
(24) Wolfenden, R. *Biochemistry* **1978**, *17*, 201.

Table I. ΔE_g Values for Energy-Refined Thyroxine Analogue-Prealbumin Complexes

	$\epsilon = 2$				$\epsilon = R_{ij}$				
	E_F[a]	E_C[b]	ΔE_g[c]	$\Delta(\Delta E_g)$[d]	E_F[a]	E_C[b]	ΔE_g[c]	$\Delta(\Delta E_g)$[d]	$\Delta(\Delta G_{exptl})$[e]
TBPA	−24.9				−86.5				
L-T_4	−13.4	−109.8	−71.5	0.0	−4.8	−172.2	−80.9	0.0	0.0
D-T_4	−13.4	−110.6	−72.3	−0.8	−4.8	−169.8	−78.5	2.4	2.01
de-NH_3^+-T_4	3.8	−92.5	−71.4	0.1	3.5	−159.2	−76.2	4.7	−0.67
de-COO^--T_4	13.1	−76.4	−64.6	6.9	10.3	−143.2	−67.0	13.9	3.45

[a] E_F = Final molecular mechanics refined energy for free analogue or binding site (kcal/mol) (rms deviation between initial and energy refined coordinates for native TBPA = 0.55 Å for both dielectric models). [b] E_C = Final molecular mechanics refined energy for analogue-binding site complex (kcal/mol). [c] $\Delta E_g = E_C - E_{F_{TBPA}} - E_{F_{analogue}}$. [d] $\Delta(\Delta E_g) = \Delta E_{g_{analogue-TBPA}} - \Delta E_{g_{L-T_4-TBPA}}$. [e] $\Delta(\Delta G_{exptl})$ = Experimental relative free energies of binding (ref 15, kcal/mol).

estimated by Wolfenden to be $\sim 1.25 \times 10^9$ by using bond contributions based on correlations of data from the literature.[25] The free energy required to transfer acetic acid, methylamine, and glycine from the gas phase to aqueous solution at pH 7.0 is $-RT \ln (K_{eq}/\alpha)$, where K_{eq} is the water/vapor distribution coefficient and α is the fraction of solute that is not ionized in aqueous solution at pH 7.0 (acetic acid,[26] 5.6×10^{-3}; methylamine,[26] 2.2×10^{-4}; glycine,[27] 5×10^{-6}). These free energies are −9.8 kcal/mol for acetic acid, −9.8 kcal/mol for methylamine, and −19.6 kcal/mol for glycine. By combining these values for $\Delta(\Delta G_{SA})$ with the values for $\Delta(\Delta E_g)$ derived from the molecular mechanics calculations, we estimate $\Delta(\Delta G_{aq})$ for the four analogues.

Initially, we performed several calculations on T_4 itself, examining the conformational profile for rotation of the phenolic C–O bond, the diphenyl ether bonds φ_1 and φ_2, and χ_1 (Figure 2). These calculations led us to represent the aromatic hydrogens of T_4 explicitly since the hydrogens ortho to the diphenyl ether oxygen were essential for the correct representation of the φ_1–φ_2 conformational profile. The set of parameters reported here led to a rotational barrier for the diphenyl ether angle φ_2 of 7.8 kcal/mol, in good agreement with values obtained from NMR data[28] (7.9 kcal/mol) and quantum mechanical calculations[29] (10.6 kcal/mol), with the minimium-energy diphenyl ether torsional angles φ_1 and φ_2 at 90 and 0°, respectively, in reasonable agreement with X-ray crystallographic results.[30] The calculated barriers for rotation around the phenolic C–O bond and χ_1 were 3.8 and 2.2 kcal/mol, respectively, with the minimium-energy conformation for the phenolic O–H bond in the ring plane and the CB–CA bond perpendicular to it. The minimium-energy diphenyl ether bond angle is 125°, somewhat larger than the value of ~120° determined by X-ray crystallography.[30]

Because the PDP-11/70 version of AMBER has a limit of 285 atoms we sought a model of the hormone-binding site of TBPA consistent with this limitation. Initially, we included all TBPA residues within 5 Å of any atom of T_4 in the binding site (13A–17A, 54A, 108A–110A, 117A–121A, 13C–17C, 54C, 108C–110C, 117C–121C), replacing the side chains of those residues extending away from the binding site with a hydrogen atom. The crystallographically well-defined bound water molecule in the site[13,14] was also included in the model. In initial energy refinements, the side chains of Glu-54A and Glu-54C (one of which interacts with the NH_3^+ group of T_4) moved up to 120° away from their original location in the X-ray structure. Although residues 50A–53A, 55A–56A, 50C–53C, and 55C–56C do not contact T_4, they are in close contact with Glu-54A and Glu-54C. In particular, His-56A and His-56C are in good position in the original X-ray structure to hydrogen bond to Glu-54A and Glu-54C, thereby greatly restricting the motion of these residues. A more realistic model of the binding site was made by adding these new residues to the binding site and removing several residues surrounding the phenolic ring of T_4 (109A–110A, 117A–120A, 109C–110C, 117C–120C) to stay within the 285-atom limit, since the phenolic ring of T_4 moved to the same position in the binding site in all initial refinements. The phenolic ring of T_4 was then restrained to this position in all subsequent refinements.

Because the X-ray evidence suggests no large-scale motion of the TBPA backbone upon T_4 binding, we restrained the peptide backbone nitrogens to their initial positions with a penalty function ($E_{restraint} = \sum_{atoms} 100D^2$, where D is the distance between the current and initial Cartesian coordinates of each atom) of 100 kcal/(mol Å^{-2}). This approach provides a reasonable representation of small sections of proteins without explicit inclusion of the entire protein molecule, while still allowing limited motion of the peptide backbone (since the CA, C, and O of the peptide backbone are unrestrained) and complete freedom for the amino acid side chains.

Initial calculations used the unmodified X-ray coordinates for native TBPA (refined at 1.8-Å resolution)[13,14] and the best fit of the T_4 iodines to the observed difference electron density map (also at 1.8-Å resolution)[31] for these groups. The latter determines only the location of the diphenyl ether rings in the hormone binding site and provides no information on the conformation of the T_4 amino acid side chain (χ_1, χ_2). The original structure had the two Lys-15 side chains in very close contact with the nonphenolic ring of T_4, so it is clear that these residues must move significantly from their original positions in native TBPA. Therefore we carried out the initial refinements in two stages: in stage one only the T_4 amino acid moiety, the two Lys-15 residues, and the two Glu-54 residues were allowed to move while keeping the remainder of the system fixed with a restraint of 1000 kcal/(mol Å^{-2}). This was done to relieve the bad contacts between the two Lys-15 residues and T_4 in the X-ray model without allowing the large energy gradients of these atoms to cause artificial movement elsewhere. In stage two the entire system was allowed to energy-refine while retaining restraints only on the motion of the peptide backbone nitrogens (100 kcal/(mol Å^{-2})). For both L- and D-T_4 the torsional angle χ_1 was then incremented systematically by 60, 120, 180, 240, and 300°, and the entire two-stage process was repeated.

Next, the energy-refined structures were each modified by real-time interactive computer graphics modeling[32] on an Evans and Sutherland Color Picture System 2 using the program CHEM.[33] This program is interfaced with the AMBER software package and allows the user to change the location of different molecules relative to each other, adjust torsional angles, and monitor interatomic distances while displaying the molecules in color and stereo. The torsional angles χ_1 and χ_2 of L- and D-T_4 and the two Lys-15 and Glu-54 side chains were adjusted to further optimize binding contacts, followed by energy refinement of these structures. This type of interactive refinement proved to be essential in finding the lowest energy structures for the TBPA-T_4 analogue complexes. We also studied the deamino and decarboxy T_4 analogues (Figure

(25) Wolfenden, R.; Andersson, L.; Cullis, *P. M.*; Southgate, C. C. B. *Biochemistry* **1981**, *20*, 849.
(26) Weasty, R. C.,; Ed. "CRCC Handbook", 56th ed.; CRC Press: Cleveland, OH, 1976.
(27) Edsall, J. T.; Wyman, J. "Biophysical Chemistry"; Academic Press: New York, 1958; p 485.
(28) Emmett, J. C.; Pepper, E. S. *Nature (London)* **1975**, *257*, 334.
(29) Andrea, T. A.; Dietrich, S. W.; Murray, W. J.; Kollman, P.; Jorgensen, E. C.; Rothenberg, S. *J. Med. Chem.* **1979**, *22*, 221.
(30) Cody, V. *Endocr. Rev.* **1980**, *1*, 140.
(31) Oatley, S. J.; Burridge, J. M.; Blake, C. C. F., unpublished results.
(32) Langridge, R.; Ferrin, T. E.; Kuntz, I. D.; Connolly, M. L. *Science (Washington, D.C.)* **1981**, *211*, 661.
(33) Dearing, A., documentation available on request.

2) with this approach, modeling them by mimicking as closely as possible the favorable ionic interactions of T_4–TBPA. From 10 to 15 different starting conformations were evaluated for each protein–ligand complex to scan as much conformational space as possible.

Results

Table I has the calculated ΔE_g ("gas phase") values for the final set of molecular mechanics energy refinements, the total energies of the isolated fragments, and the experimental free energies of binding for the T_4 analogue–TBPA interactions. With $\epsilon = R_{ij}$, both the difference in binding affinity between L- and D-T_4 and the difference in binding affinity between deamino and decarboxy T_4 are reasonably well reproduced in the calculations, whereas in the calculations with $\epsilon = 2.0$, only the deamino–decarboxy difference is qualitatively reproduced, while the relative L- and D-T_4 binding affinities are reversed. The amino acids (L- and D-T_4) are calculated to bind more strongly relative to the carboxylic acid (deamino) or amine (decarboxy) analogues than found experimentally. In Table II we summarize our estimated $\Delta(\Delta G_{aq})$ ("corrected" for differential solvation effects) values for $\epsilon = R_{ij}$ and note that they are now in qualitative agreement with the experimental results. The agreement with $\epsilon = 2$ is poorer.

Although the calculated ΔE_g values for the four analogue–TBPA complexes are sensitive to ϵ, the *structural* results are consistent and independent of ϵ, with nearly identical structures at each ϵ. The rms deviation between the $\epsilon = R_{ij}$ structure and the $\epsilon = 2$ structure is 0.20–0.28 Å for the four analogue–TBPA complexes.

After the initial bad contacts of the two Lys-15 side chains were relieved with the hormone, energy refinement with the original binding site model led to a structure in which T_4 had moved ~0.7 Å deeper into the binding site, in close agreement with the most recent interpretation of the difference electron density map for T_4[34] these coordinates were not available at the beginning of our study). In all refinements, the diphenyl ether torsional angles ϕ_1 and ϕ_2 remained near their minimum energy values of 90 and 0°. The bound water molecule refined to a position between Ser-117A and Thr-119A in all refinements, forming hydrogen bonds with both residues (the H_2O–O···Ser-117A–OH distance = 1.91 Å, while one of the H_2O hydrogens contacts the Thr-119A hydroxyl oxygen at 1.95 Å). The water apparently favors a tightly bound location between these two residues rather than hydrogen bonding to the phenolic hydroxyl group of T_4. The phenolic hydroxyl group of the hormone is too far from the water, Ser-117A, or Thr-119A to form a strong hydrogen bond and does not appear to have any significant interaction with the TBPA binding site, consistent with recent studies comparing the TBPA binding affinities of 4′-OH-T_4 analogues with 4′-H-T_4 analogues that demonstrate that the phenolic hydroxyl group contributes very little to the free energy of binding.[35]

The lowest energy L- and D-T_4–TBPA complexes are structurally very similar (Figure 3). In both models, the T_4 amino acid moiety spans across the binding site from the A subunit to the C subunit, with the NH_3^+ interacting with Glu-54A and the COO^- with Lys-15C. These "bridging" structures were 4–5 kcal lower in energy than the lowest energy "nonbridging" structures (where the NH_3^+ and COO^- groups of T_4 interact with only one subunit). Both L- and D-T_4 form an extensive ion-pairing network with the binding site of TBPA as follows: His-56A–HNE···$^-$OOC–Glu-54A; Glu-54A–COO^-···^+H_3N–Lys-15A; Lys-15A–NH_3^+···O–Ser-52A; Glu-54A–COO^-···^+H_3N–T_4; T_4–COO^-···^+H_3N–Lys-15C; Lys-15C–NH_3^+···$^-$OOC–Glu-54C; Glu-54C–COO^-···HNE–His-56A. The ion-pairing (H--O) distances range from 1.34 to 1.94 Å with $\epsilon = R_{ij}$ and from 1.60 to 2.01 Å with $\epsilon = 2$. The stereoselectivity in binding to TBPA for L-T_4 over D-T_4 apparently involves a delicate balance between opposing forces and must result from a complex interplay of ionic interactions rather than steric effects. From Figure 4 ($\epsilon = R_{ij}$), the major

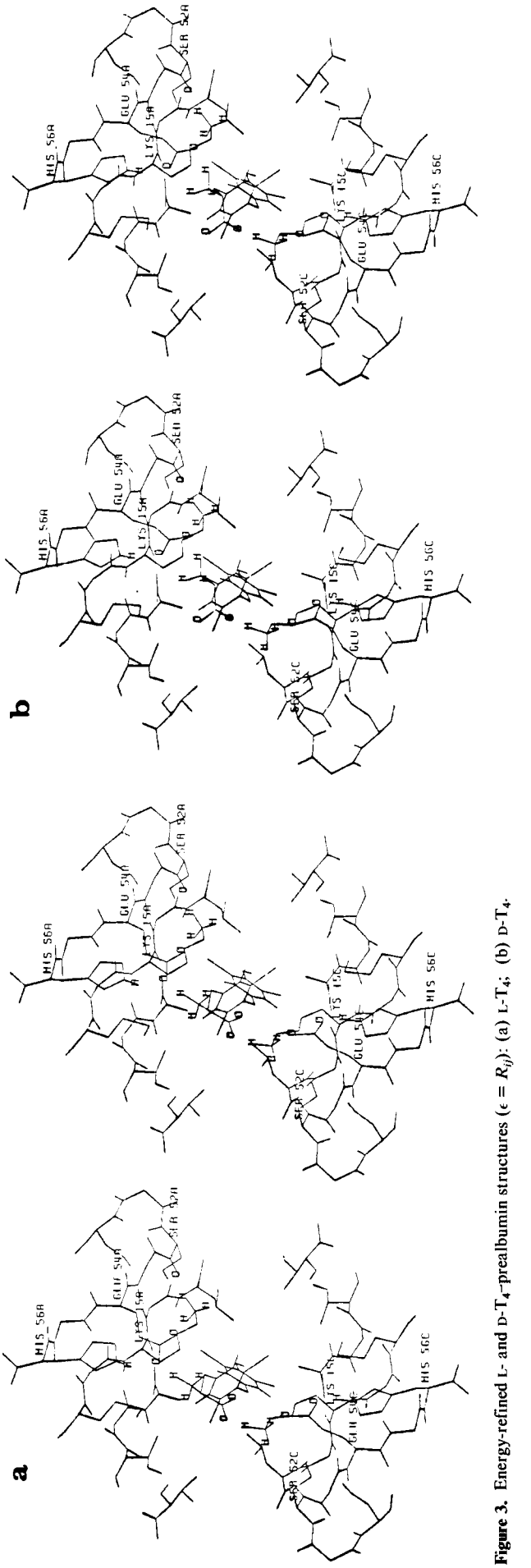


Figure 3. Energy-refined L- and D-T_4–prealbumin structures ($\epsilon = R_{ij}$): (a) L-T_4; (b) D-T_4.

(34) Burridge, J. M.; Oatley, S. J.; Blake, C. C. F., unpublished results.
(35) Dearing, A.; Weiner, P.; Kollman, P. *Nucleic Acids Res.* **1981**, *9*, 1483.

Table II. Solvation Model for Energy-Refined Thyroxine Analogue–Prealbumin Complexes

		$\epsilon = 2$		$\epsilon = R_{ij}$		
	$\Delta(\Delta G_{SA})^a$	$\Delta(\Delta E_g)^b$	$\Delta(\Delta G_{aq})^c$	$\Delta(\Delta E_g)^b$	$\Delta(\Delta G_{aq})^c$	$\Delta(\Delta G_{exptl})^d$
L-T_4	−19.8	0.0	0.0	0.0	0.0	0.0
D-T_4	−19.6	−0.8	−0.8	2.4	2.4	2.01
de-NH_3^+-T_4	−9.8	0.1	−9.7	4.7	−5.1	−0.67
de-COO^--T_4	−9.8	6.9	−2.9	13.9	4.1	3.45

[a] $\Delta(\Delta G_{SA})$ = relative free energy of solvation of analogue (kcal/mol). [b] $\Delta(\Delta E_g)$ = relative free energy of binding of analogue (gas phase, kcal/mol). [c] $\Delta(\Delta E_{aq})$ = relative free energy of binding of analogue (aqueous, kcal/mol, values adjusted so $\Delta(\Delta G_{aq})_{L\text{-}T_4} = 0.0$). [d] $\Delta(\Delta G_{exptl})$ = experimental relative free energies of binding (ref 15, kcal/mol).

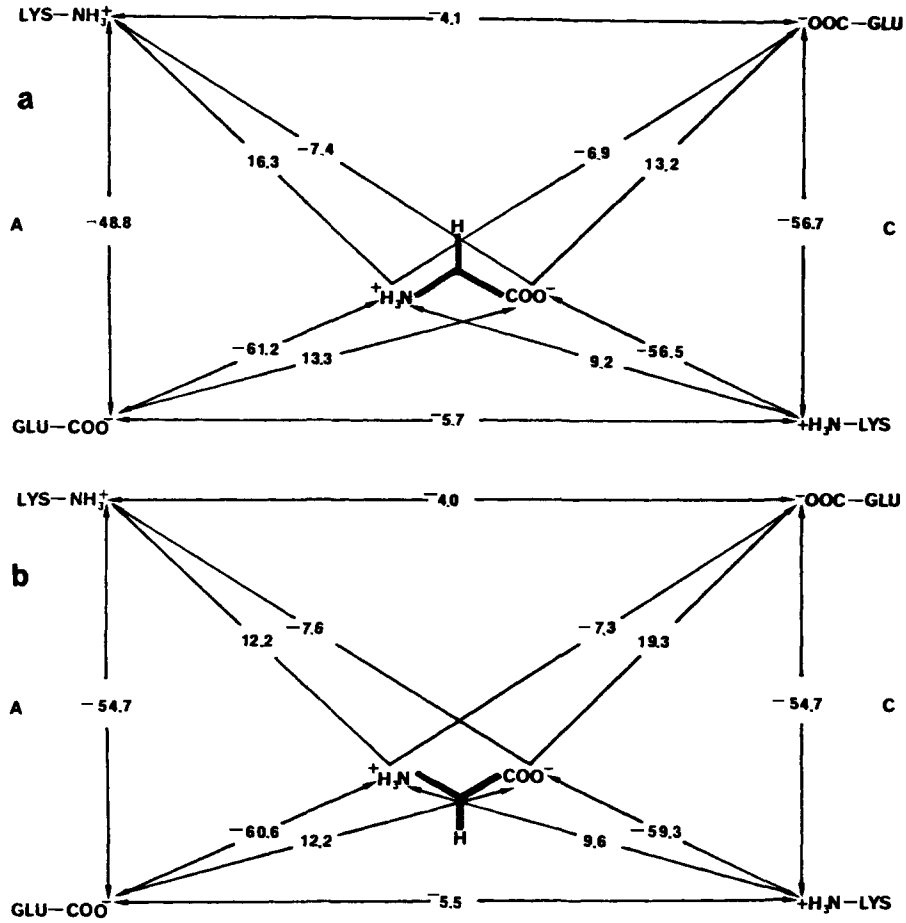


Figure 4. Interaction energies (kcal/mol) for $\epsilon = R_{ij}$ model: (a) L-T_4; (b) D-T_4.

factors responsible for the L/D stereoselectivity are the T_4–$COO^-\cdots{}^+H_3N$–Lys-15C, the T_4–$NH_3^+\cdots{}^+H_3N$–Lys-15A, and the Lys-15A–$NH_3^+\cdots{}^-OOC$–Glu-54A interactions, which favor the L enantiomer by 2.8, 4.1, and 5.9 kcal/mol, respectively. The interactions between the Glu-54C–$COO^-\cdots{}^+H_3N$–Lys-15C and the T_4–$COO^-\cdots{}^-OOC$–Glu-54C favor the D enantiomer by 2.0 and 6.1 kcal/mol, respectively. The remaining small differential interactions combine to favor the D enantiomer by 2.3 kcal/mol, leading to the net result (Tables I and II) that the L enantiomer is correctly calculated to be more tightly bound to the protein by 2.4 kcal/mol with $\epsilon = R_{ij}$.

The lowest energy deamino and decarboxy-T_4–TBPA structures are significantly different and immediately suggest why deamino-T_4 has a greater binding affinity than decarboxy-T_4. The deamino-T_4–COO^- group makes two favorable ionic interactions by contacting the NH_3^+ groups of *both* Lys-15A and Lys-15C near the center of the binding site (Figure 5a), while a similar interaction of the decarboxy-T_4–NH_3^+ group with the COO^- groups of Glu-54A and Glu-54C is impossible (Figure 5b); the glutamate side chains are shorter than the lysine side chains and are therefore unable to reach the center of the binding site, so the decarboxy analogue is only able to make one ionic interaction with a single glutamate (Glu-54A). In the deamino structure, the Lys-15A–NH_3^+ has rotated to the opposite side of the Glu-54A–COO^- in order to interact better with the deamino-T_4–COO^- group. This structure was 6–7 kcal more stable than the structure with the Lys-15A side chain in its original position in the X-ray structure. Comparison of the group–group interaction energies in Figure 6a,b ($\epsilon = R_{ij}$) shows that the difference in calculated interaction energies between the charged groups of the protein with the decarboxy and deamino analogues favors deamino-T_4 by ~40 kcal/mol. However, in the decarboxy-T_4 complex, the Lys-15A and Lys-15C side chains make favorable contacts with the rest of the protein (not shown in Figure 6), in particular forming hydrogen bonds with the peptide backbone carbonyl oxygens of Ser-52A and Ser-52C (Figure 5b), favoring the decarboxy analogue by ~30 kcal/mol so that the net difference in the overall calculated binding energies for the deamino and the decarboxy analogues is ~10 kcal/mol with $\epsilon = R_{ij}$.

Discussion

The amino acid analogues L- and D-T_4 are calculated ("gas phase", Table I) to interact more strongly with the TBPA binding site than either deamino or decarboxy-T_4, regardless of the choice of ϵ. This is not surprising since the hormone binding site can form several more attractive ionic interactions with the zwitterionic

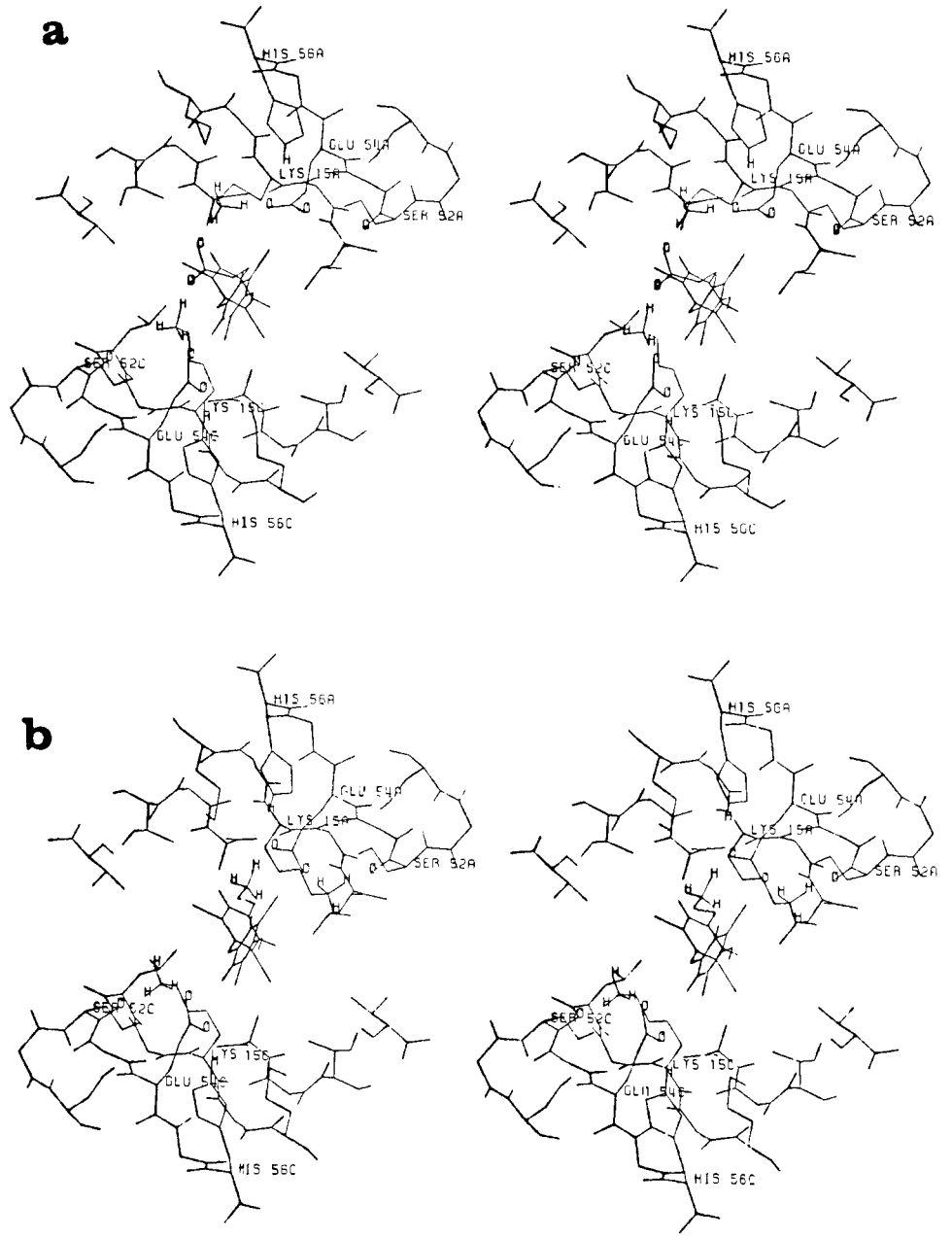


Figure 5. Energy-refined deamino- and decarboxy-T_4–prealbumin structures ($\epsilon = R_{ij}$): (a) deamino-T_4; (b) decarboxy-T_4.

amino acid analogues than with the singly charged analogues. So why is deamino-T_4 experimentally observed to be more tightly bound to TBPA than L- or D-T_4? Desolvation of the amino acid analogues upon binding to TBPA costs more free energy than desolvation of the deamino or decarboxy analogues. Therefore, deamino-T_4 is more tightly bound to TBPA because its desolvation costs ~10 kcal/mol less free energy than the desolvation of L- and D-T_4, *not* because of intrinsically stronger interaction with the TBPA binding site (Table II).

The buried surface areas (the solvent-accessible surface area of the unbound protein and hormone minus the solvent-accessible surface area when complexed) of the charged protein and hormone groups (NH_3^+ and COO^-) for each energy-refined complex are 33, 38, and 31, and 18 Å^2 for **1**–**4**, respectively. The approximation that $\Delta(\Delta G_{SC}) = 0$ thus underestimates the solvation energy for complex formation with **4**; we were unable to find any method for converting these differences in solvent-accessible surface areas into solvation energies that gave a better agreement with the experimental values for $\Delta(\Delta G_{expt})$.

The major difference between the $\epsilon = 2.0$ and $\epsilon = R_{ij}$ $\Delta(\Delta E)$ results is that the charged analogues **3** and **4** are less stabilized relative to **1** and **2** for the $\epsilon = R_{ij}$ model, since the interaction energies with $\epsilon = R_{ij}$ are more dominated by the attractive close-contact interactions. This is reflected in the shorter ion-pairing distances found with $\epsilon = R_{ij}$ relative to $\epsilon = 2$. Considering the hydrogen-bonding networks shown in Figure 3 and 5, the simplest representations of the TBPA–analogue complexes are as one-dimensional ionic lattices containing six charges (+−+−+−) for L- and D-T_4 and five charges for deamino-T_4 (−+−+−) and decarboxy-T_4 (+−+−+). With the assumption of equal distances between charges, calculating the electrostatic energy for these one-dimensional lattices leads to a lower energy for L- and D-T_4 ($\Delta E = \sum_{i=1}^{5}(-1)^i/\epsilon iR$, where ΔE is the difference between the six and five lattice point energies, ϵ is the dielectric constant, and R is the distance between neighboring lattice points) than for deamino-T_4 or decarboxy-T_4. If we let $\epsilon = 2$ and $R = 1$, $\Delta E = -{}^1/_2 + {}^1/_4 - {}^1/_6 + {}^1/_8 - {}^1/_{10} = -0.39$. However, if we use a distance-dependent dielectric model where $\epsilon = iR$, $\Delta E = -1 + {}^1/_4 - {}^1/_9 + {}^1/_{16} - {}^1/_{25} = -0.83$. Thus, the amino acids L- and D-T_4 are more stabilized relative to the amine or carboxylic acid analogues when a distance-dependent dielectric model is used (Table I).

Our previous much simpler analysis of T_4 binding to TBPA concluded that steric interactions of the T_4–amino acid moiety

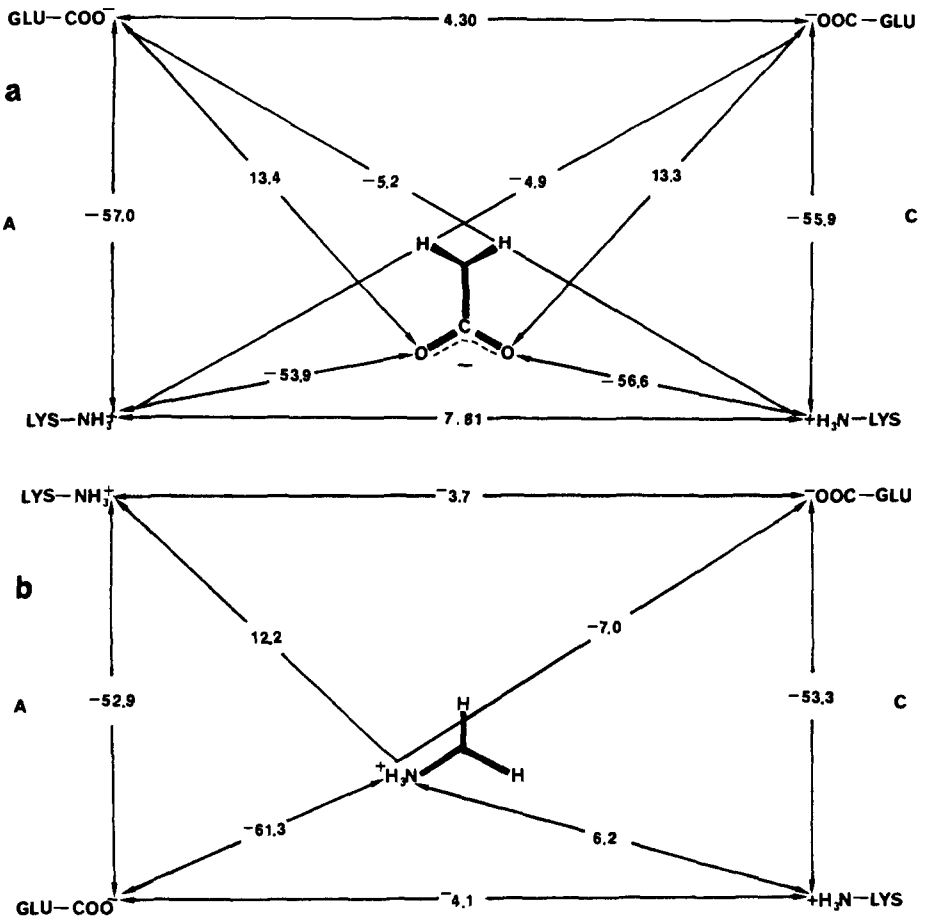


Figure 6. Interaction energies (kcal/mol) for $\epsilon = R_{ij}$ model: (a) deamino-T_4; (b) decarboxy-T_4.

with Leu-17 and Val-121 were primarily responsible for the L/D stereoselectivity.[36] However, this early work used a simple hard-sphere representation for each amino acid in the hormone-binding site and was based on preliminary coordinates for the native protein that did not include the bound water molecule or T_4. The modeling reported here uses more recent X-ray data for the T_4–TBPA complex and a much more sophisticated all-atom molecular mechanical model and should provide a more realistic picture of the binding interaction. In fact, Leu-17 and Val-121 are too far away from the T_4 amino acid moiety to influence the binding stereoselectivity. We now find that the L/D stereoselectivity in binding to TBPA must result primarily from a complex interplay of ionic interactions rather than steric effects.

The calculations with $\epsilon = R_{ij}$ including a simple solvation model are the best in reproducing the experimental binding data for the thyroid hormone analogues. A distance-dependent dielectric constant has been reasonably successful in other molecular mechanics studies of nucleic acids and proteins[18,37-38] and suggests that, with appropriate solvation models such as employed here, this method might be generally useful in studying the intermolecular interactions in complex systems.

Given that this is the first time such models and molecular mechanics have been used together to compare the noncovalent association of anionic, neutral, and cationic ligands to a macromolecule, the qualitative agreement with experiment with use of $\epsilon = R_{ij}$ and empirical solvation corrections is encouraging. However, due to the crudeness of our model and the fact that we are only able to explore a small amount of conformational space for detailed energy analysis, we must emphasize that it is still quite difficult to predict reliably (rather than reproduce) the relative strengths of most ligand–macromolecule interactions.

(36) Somack, R.; Andrea, T. A.; Jorgensen, E. C. *Biochemistry* **1982**, *21*, 163.
(37) Andrea, T. A.; Jorgensen, E. C.; Kollman, P. *Int. J. Quantum Chem., Quantum Biol. Symp.* **1978**, *No. 5*, 191.
(38) Wipff, G.; Dearing, Weiner, P.; Blaney, J.; Kollman, P. *J. Am. Chem. Soc.*, in press.

Summary

We have simulated the interaction of L-T_4 and three analogues with a model of the thyroid hormone binding site of the plasma protein TBPA using molecular mechanics calculations. The molecular mechanics calculations alone proved sufficient to reproduce the relative free energies of binding for the charged analogues deamino-T_4 (**3**) and decarboxy-T_4 (**4**) due to the ability of both lysine NH_3^+ groups in the binding site to reach and ion pair with the COO^- group of **3**, whereas only one of the two binding-site glutamate side chains can interact with the NH_3^+ group of **4**. Correctly modeling the stereoselectivity in binding of L-T_4 over D-T_4 was much more difficult; there is no single dominant (e.g., steric) ligand–protein interaction that discriminates between the two enantiomers. Instead, the stereoselectivity results from a large number of contributing factors, involving charge–charge interactions between the zwitterionic amino acid portion of the hormone and the two lysine and two glutamate residues in the binding site. To reproduce the experimental relative free energies of binding for all four analogues required the inclusion of solvation effects, which we have modeled by using a simple empirical method. Using the experimental free energies of transfer of acetic acid and methylamine and an estimate of the free energy of transfer for glycine from the gas to the aqueous phase, we have estimated the relative solvation energies of the four thyroid hormone analogues and successfully reproduced the order of binding affinities for the analogues. Although the amino acids L- and D-T_4 form more attractive ionic interactions and therefore interact more strongly with TBPA than deamino or decarboxy-T_4, desolvation of the zwitterionic amino acid analogues upon binding to TBPA costs more free energy than desolvation of the singly charged deamino or decarboxy analogs. Therefore, deamino-T_4 is more tightly bound to TBPA than L- or D-T_4 *not* because of intrinsically stronger interaction with the TBPA binding site but

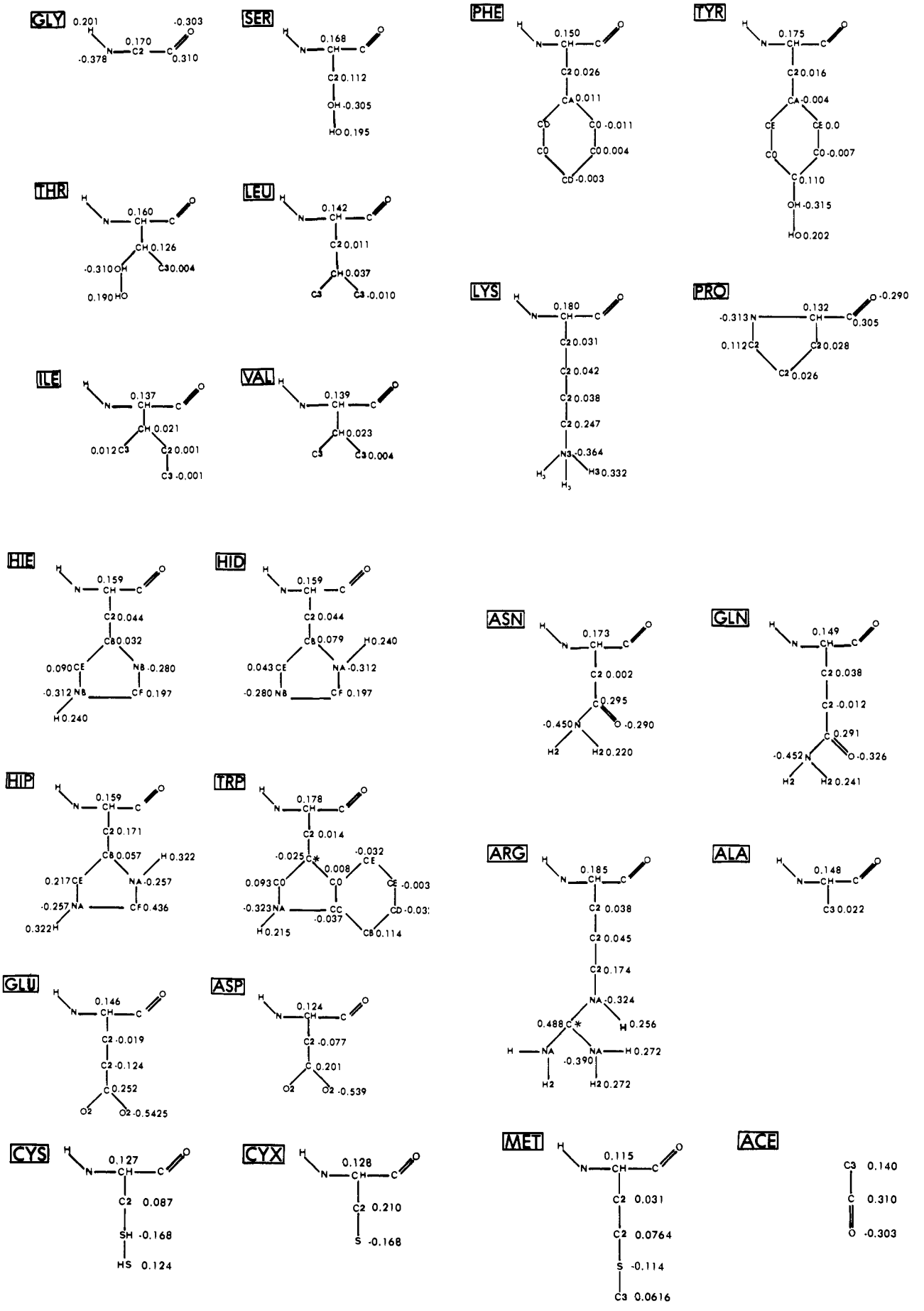


Figure 7. Atomic partial charges: amino acid charges (the backbone charges are identical with those given for Gly).

Table III. Potential Function Parameters. See Figures 7 and 8 for the Atom Types of the Various Atoms

I. Peptide Parameters

bond stretching parameters

A1-A2	K_R	R_0	A1-A2	K_R	R_0	A1-A2	K_R	R_0	A1-A2	K_R	R_0
C-C3	300	1.52	CF-C*	450	1.34	CB-CD	450	1.39	C*-NA	450	1.33
C-CB	450	1.42	CF-NB	450	1.36	CB-C*	450	1.42	H-NA	300	1.01
C-CF	450	1.44	CH-N	300	1.45	CB-CB	450	1.37	H-N	300	1.01
C-NA	450	1.39	CH-C2	300	1.52	CB-NB	450	1.39	H2-N	300	1.01
C-O	600	1.23	CH-C3	300	1.52	CB-C2	450	1.51	HO-OH	300	0.96
C-OH	450	1.36	CH-CH	300	1.52	CB-CE	450	1.36	HS-SH	300	1.33
C-N	450	1.34	CH-OH	300	1.43	CB-NA	450	1.39	H3-N3	300	1.01
C-CD	450	1.44	C2-C2	300	1.52	CC-CD	450	1.39	S-S	300	2.01
C-C2	300	1.52	C2-OH	300	1.43	CC-NA	450	1.43	CH-CA	300	1.51
C-O2	450	1.25	C2-NA	300	1.47	CD-CE	450	1.39	CA-C	450	1.44
C-CH	300	1.52	C2-C*	300	1.51	CD-NA	450	1.43	N-C3	300	1.47
CA-CB	450	1.40	C2-N	300	1.47	CD-C*	450	1.35	C-CA	450	1.40
CA-CF	450	1.43	C2-C3	300	1.52	CD-CD	450	1.39	CA-C3	300	1.51
CA-NA	450	1.38	C2-SH	300	1.80	CE-NA	450	1.36	CD-C3	300	1.51
CA-C2	300	1.51	C2-S	300	1.80	CE-NB	450	1.30	CA-CA	450	1.40
CA-CD	450	1.39	C2-N3	300	1.47	CE-CE	450	1.39	H-CD	300	1.08
CA-CE	450	1.39	C3-C*	300	1.52	CF-NA	450	1.32	H-CE	300	1.08
CB-CC	450	1.39	C3-S	300	1.80	CF-CF	450	1.35			

bond bending parameters

A1-A2-A3	K_θ	θ_0	A1-A2-A3	K_θ	θ_0	A1-A2-A3	K_θ	θ_0	A1-A2-A3	K_θ	θ_0
X-C2-X	46.5	112.0	CH-CH-CH	46.5	111.0	C2-CA-CE	70.0	120.0	CD-C*-C2	70.0	127.0
X-C-X	70.0	120.0	C2-CH-N	46.5	111.0	CE-CA-CE	70.0	120.0	CB-C*-CD	70.0	108.0
X-CA-X	70.0	120.0	C-CH-N	46.5	111.0	C2-CB-NA	70.0	122.0	C-N-H	70.0	119.8
X-CB-X	70.0	120.0	C-CH-C2	46.5	111.0	C2-CB-CE	70.0	132.0	C-N-CH	70.0	121.5
X-CC-X	70.0	120.0	N-CH-C2	46.5	111.0	CE-CB-NA	70.0	106.0	CH-N-H	70.0	118.3
X-CD-X	70.0	120.0	C-CH-C3	46.5	111.0	CB-CB-NB	70.0	110.4	C-N-C2	70.0	121.5
X-CE-X	70.0	120.0	N-CH-CH	46.5	111.0	C-CB-NB	70.0	130.0	CH-N-C2	70.0	112.2
X-CF-X	70.0	120.0	CH-CH-C3	46.5	111.0	C-CB-CB	70.0	119.2	C2-N-H	70.0	118.3
X-NB-X	70.0	120.0	C2-CH-C3	46.5	111.0	CA-CB-NB	70.0	132.4	CF-NA-H	70.0	126.0
X-N-X	70.0	120.0	CH-CH-OH	46.5	111.0	CA-CB-CB	70.0	117.3	CB-NA-CF	70.0	108.0
X-NA-X	70.0	120.0	C3-CH-OH	46.5	111.0	C2-CB-NB	70.0	122.0	CB-NA-H	70.0	126.0
X-N3-X	46.5	109.5	C3-CH-C3	46.5	111.0	CE-CB-NB	70.0	106.0	CE-NA-H	70.0	125.0
O-C-OH	70.0	120.0	C-CH-CH	46.5	111.0	CC-CB-C*	70.0	108.0	CE-NA-CF	70.0	110.0
C3-C-N	70.0	120.0	CH-C2-CH	46.5	112.0	CD-CB-C*	70.0	132.0	CC-NA-H	70.0	125.5
C3-C-O	70.0	120.0	CB-C2-CH	46.5	115.0	CC-CB-CD	70.0	120.0	CC-NA-CD	70.0	109.0
CD-C-OH	70.0	120.0	CH-C2-C*	46.5	115.0	CD-CC-NA	70.0	128.0	CD-NA-H	70.0	125.5
CD-C-CD	70.0	120.0	C2-C2-N3	46.5	112.0	CB-CC-CD	70.0	120.0	C2-NA-H	70.0	118.5
CF-C-O	70.0	125.3	C-C2-C2	46.5	112.0	CB-CC-NA	70.0	107.0	C*-NA-H	70.0	120.0
NA-C-O	70.0	120.6	C-C2-CH	46.5	112.0	CA-CD-CD	70.0	120.0	C2-NA-C*	70.0	120.0
CF-C-NA	70.0	114.1	CH-C2-OH	46.5	112.0	CD-CD-CD	70.0	120.0	C-NA-H	70.0	116.5
C*-C-O	70.0	125.3	C2-C2-NA	46.5	112.0	C*-CD-NA	70.0	107.0	CA-NA-H	70.0	119.3
C*-C-NA	70.0	114.1	C-C2-N	46.5	112.0	CC-CD-CE	70.0	120.0	C-NA-CA	70.0	125.2
CB-C-O	70.0	128.8	CH-C2-C3	46.5	112.0	CB-CD-CE	70.0	120.0	C-NA-C	70.0	126.4
CB-C-NA	70.0	111.3	CH-C2-SH	46.5	115.7	CE-CD-C	70.0	120.0	CE-NB-CF	70.0	110.0
CH-C-O2	70.0	120.4	CH-C2-S	46.5	115.7	CB-CE-NB	70.0	107.0	CB-NB-CE	70.0	103.8
CH-C-O	70.0	120.4	CH-C2-C2	46.5	112.0	CB-CE-NA	70.0	107.0	C2-N3-H3	46.5	109.5
N-C-O	70.0	123.5	C2-C2-S	46.5	115.7	CD-CE-CE	70.0	120.0	H3-N3-H3	46.5	109.5
CH-C-N	70.0	115.7	C2-C2-N	46.5	112.0	CA-CE-CD	70.0	120.0	CH-OH-HO	46.5	107.3
C2-C-N	70.0	115.7	C2-C2-C2	46.5	112.0	NB-CF-NA	70.0	109.0	C-OH-HO	46.5	113.0
C2-C-O2	70.0	120.4	CH-C2-CA	46.5	115.0	CA-CF-CF	70.0	117.0	C2-OH-HO	46.5	107.3
O2-C-O2	70.0	120.0	C-C*-CF	70.0	120.7	C-CF-CF	70.0	120.7	C2-S-C3	46.5	104.0
C2-C-O	70.0	120.4	C-C*-C3	70.0	119.7	NA-CF-NA	70.0	109.0	C2-S-S	46.5	104.0
C2-CA-CD	70.0	120.0	CF-C*-C3	70.0	119.7	C2-CH-OH	46.5	111.0	C2-SH-HS	46.5	100.0
CD-CA-CD	70.0	120.0	CB-C*-C2	70.0	125.0	CH-CH-C2	46.5	111.0	C2-CH-CA	46.5	115.0
									HO-OH-HO	46.5	105.0

torsional parameters

A1-A2-A3-A4	no. of bonds	K_N	δ	N	A1-A2-A3-A4	no. of bonds	K_N	δ	N
X-CH-CH-X	4	1.5	0	3	X-C-NA-X	4	25.0	180	2
X-CF-CF-X	1	25.0	180	2	X-NB-CB-X	2	25.0	180	2
X-CF-C*-X	2	25.0	180	2	X-C-CB-X	4	25.0	180	2
X-CA-CF-X	2	25.0	180	2	X-NA-CA-X	4	25.0	180	2
X-C-C*-X	4	25.0	180	2	X-CF-NB-X	1	25.0	180	2
X-CD-CD-X	1	25.0	180	2	X-CA-CB-X	4	25.0	180	2
X-CB-CB-X	4	25.0	180	2	X-CE-NB-X	1	25.0	180	2
X-CD-CA-X	2	25.0	180	2	X-S-S-X	1	6.0	0	2
X-CE-CA-X	2	25.0	180	2	X-S-C2-X	1	1.0	0	3
X-CA-NB-X	2	25.0	180	2	X-N-CH-X	4	0.6	0	3
X-CH-OH-X	2	0.5	0	3	X-CH-C-X	4	0.7	0	3
X-CD-C-X	2	25.0	180	2	X-C-N-X	4	10.0	180	2

Table III (*Continued*)

torsional parameters

A1-A2-A3-A4	no. of bond	K_N	δ	N	A1-A2-A3-A4	no. of bonds	K_N	δ	N
X-C2-C2-X	1	1.5	0	3	X-C-OH-X	2	2.0	180	2
X-C2-NA-X	2	1.5	90	2	X-CB-CC-X	4	25.0	180	2
X-CF-NA-X	2	25.0	180	2	X-CD-NA-X	2	25.0	180	2
X-C2-C-X	2	0.7	0	6	X-CB-CD-X	2	25.0	180	2
X-C2-SH-X	1	1.0	0	3	X-NA-CC-X	4	25.0	180	2
X-C2-CB-X	2	1.5	90	2	X-CC-CD-X	2	25.0	180	2
X-CB-NA-X	4	25.0	180	2	X-CE-CE-X	1	25.0	180	2
X-CB-CE-X	2	25.0	180	2	X-C*-CB-X	4	25.0	180	2
X-CE-NA-X	2	25.0	180	2	X-C*-CD-X	4	25.0	180	2
X-C2-N3-X	3	1.5	0	3	X-C2-C*-X	2	1.5	90	2
X-C2-CA-X	2	0	0	2	X-CH-CA-X	2	1.5	90	2
X-N-C2-X	2	0.6	0	3	X-CA-C-X	4	25.0	180	2
X-C2-OH-X	1	0.5	0	3	X-CA-CA-X	4	25.0	180	2
X-CE-CD-X	1	25.0	180	2					

improper torsional parameters

A1-A2-A3-A4	K_N	δ	N	A1-A2-A3-A4	K_N	δ	N
C-X-N-NH	7.0	180	2	X-CH-N-C	7.0	180	3
X-N-C-O	7.0	180	2	C3-CH-CH-C2	7.0	180	3
X-O2-C-O2	7.0	180	2	X-CH-X-X	7.0	180	3

equivalences in van der Waals list

C	C*	CB	CC	CA
CD	CE	CF		
N	NA	N2	N*	
NB	NC			
OH	OS			
S	SH			

van der Waals terms

atom	α	N_{eff}	R	atom	α	N_{eff}	R
HO	0.42	0.9	1.30	C2	1.77	7.0	1.90
H2	0.42	0.9	1.30	C3	2.17	8.0	1.95
H3	0.42	0.9	1.30	C	1.65	5.0	1.80
HS	0.42	0.9	1.30	CD	2.07	6.0	1.90
OH	0.59	7.0	1.60	N	1.15	6.0	1.65
O	0.84	7.0	1.60	NB	1.15	6.0	1.65
O2	2.14	7.0	1.60	N3	0.87	6.0	1.65
CH	1.35	6.0	1.85	S	2.00	15.0	1.90

II. Additional Parameters Used for Thyroid Hormones

bond stretching parameters

A1-A2	K_R	R_0	A1-A2	K_R	R_0
CA-OS	300	1.40	CA-I	100	2.08
CH-N3	300	1.47	H-CA	300	1.08
CH-CA	300	1.51			

bond bending parameters

A1-A2-A3	K_θ	θ_0	A1-A2-A3	K_θ	θ_0
X-OS-X	46.5	111.0	HO-OH-HO	46.5	105.0
CH-C-O2	70.0	120.4	CA-OS-CA	46.5	120.0
CH-C2-CA	46.5	115.0			

torsional parameters

A1-A2-A3-A4	no. of bonds	K_N	δ	N
X-CH-N3-X	3	1.5	0	3
X-CA-OS-X	2	2.0	180	2

improper torsional parameters

A1-A2-A3-A4	K_N	δ	N
X-O2-C-O2	7.0	180	2
C2--CH-X-X	7.0	180	3

van der Waals terms

atom	α	N_{eff}	R
I	4.8	53.0	2.15

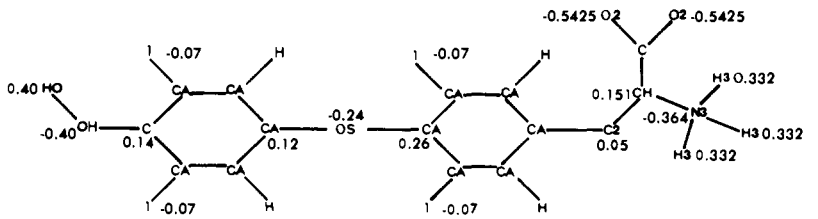


Figure 8. Charges used for thyroxine: the deamino- and decarboxy-thyroxine analogues have the same charges for the NH_3^+ and COO^-, with the charge on the chiral CH modified to create a unit charge.

because desolvation of deamino-T_4 costs about 10 kcal/mol less free energy than desolvation of L- or D-T_4.

Although TBPA is not a true drug receptor, acting only as a transport protein for T_4, the T_4-TBPA complex provides a simple working model of a drug-receptor interaction. The successful calculation of the experimental binding affinities to TBPA for analogues **1-4** illustrates a general method that may ultimately be useful for the design of new analogues and the prediction of their binding affinities.

Note Added in Proof. We have further refined the parameters and charges presented in Table III (S. J. Weines et al., to be submitted) and these are available from the authors on request.

Acknowledgment. We thank Professor R. Langridge for the use of the UCSF computer graphics laboratory, supported by Grant RR-1081. This work was supported by NIH Grants AM-17576 (E.C.J.), GM-29072 and CA-25644 (P.A.K.), the American Foundation for Pharmaceutical Education (J. M. Blaney), and the Medical Research Council of Great Britain. S.J.O. is a Mr. and Mrs. John Jaffé Donation Research Fellow of the Royal Society.

Registry No. 1, 51-48-9; **2**, 83208-10-0; **3**, 67-30-1; **4**, 83208-11-1.

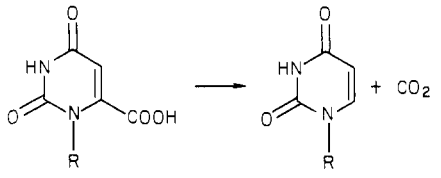
Model Chemistry for a Covalent Mechanism of Action of Orotidine 5′-Phosphate Decarboxylase

Richard B. Silverman*† and Michael P. Groziak‡

Contribution from the Department of Chemistry, Northwestern University, Evanston, Illinois 60201. *Received February 26, 1982*

Abstract: Orotidine 5′-phosphate decarboxylase (ODase) catalyzes the conversion of orotidylate to uridylate, the last step in the de novo biosynthesis of pyrimidine nucleotides. Model reactions are described that support a covalent catalytic mechanism for this enzyme in which, following protonation of the carboxyl group of orotidylic acid, an active-site nucleophile undergoes a Michael addition to the C-5 position. This covalent complex breaks down via an acid-base-catalyzed decarboxylative elimination reaction to give uridylate and CO_2 (Scheme II). The enzyme mechanism is modeled in two parts, the Michael addition reaction and the decarboxylative elimination. Bisulfite is shown to undergo a Michael addition to *N,N*-dimethylorotaldehyde and at room temperature to *N,N*-dimethyl-6-acetyluracil, both models for the activated form of orotidylate, the substrate for ODase (**6** → **7**). In a separate study, (±)-1,3-dimethyl-*r*-5-(methylthio)-5-methyl-*trans*-6-carboxyl-5,6-dihydrouracil (**15**) was prepared as a model for the ODase-orotidylate covalent complex. Activation by methylation of the sulfide (as a model for enzyme-catalyzed protonation) leads to instantaneous decarboxylative elimination at room temperature. When the corresponding ester (**9c**) is methylated, the dimethylsulfonium salt (**16b**) can be isolated, which upon ester hydrolysis gives the decarboxylative elimination product. These model studies support the Michael addition-decarboxylative elimination mechanism in favor of a noncovalent mechanism previously reported (Beak, P.; Siegel, B. *J. Am. Chem. Soc.* **1976**, *98*, 3601).

The final step in the de novo biosynthesis of pyrimidine nucleotides is the decarboxylation of orotidylic acid (**1**) to uridylic

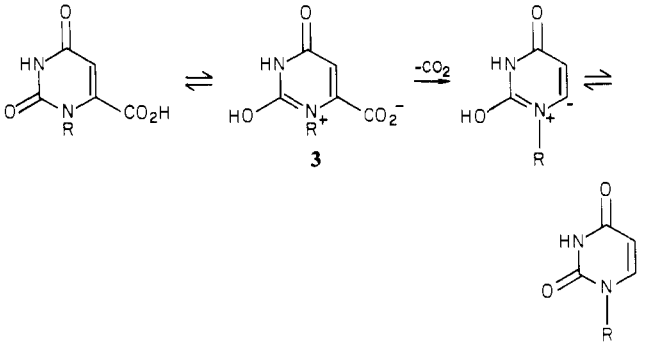


1 **2**

R = ribose 5′-phosphate

acid (**2**), a reaction catalyzed by orotidine 5′-phosphate decarboxylase (ODase; E. C. 4.1.1.23).[1] Uridylic acid can be metabolized further to cytidine and thymidine nucleotides and thus is a precursor to both RNA and DNA.[1] Malfunctioning of ODase has been shown to be responsible for the metabolic disease known as hereditary orotic aciduria.[2] The mechanism of action of this enzyme has not been delineated, but model studies[3] and enzyme inhibition studies[4] previously reported were interpreted to suggest that decarboxylation may occur from a noncovalent enzyme-stabilized zwitterion of the substrate (**3**, Scheme I) and that inductive stabilization of the resultant sp^2 carbanion is provided by the adjacent quaternary ammonium atom. Most of

Scheme I. Noncovalent Mechanism for ODase[3] (R = Ribose 5′-Phosphate)



3

the model reactions,[3] however, were carried out at very high temperatures (180-220 °C), and although the mechanism for thermal decarboxylation of 1,3-dimethylorotic acid may proceed as was suggested, it did not seem likely to us to be an efficient enzyme-catalyzed mechanism. An alternative mechanism is the

† Alfred P. Sloan Research Fellow, 1981-1983; NIH Research Career Development Awardee, 1982-1987.

‡ Present address: Department of Pharmacology, The University of Michigan, Ann Arbor, MI 48109.

(1) Hartman, S. C. In "Metabolic Pathways", 3rd ed.; Greenberg, D. M., Ed.; Academic Press: 1970; Vol. IV.

(2) Smith, L. H., Jr.; Huguley, C. M., Jr.; Bain, J. A. In "The Metabolic Basis of Inherited Disease", 3rd ed.; Stanbury, J. B., Wyngaarden, J. B., Fredrickson, D. S., Eds.; McGraw-Hill: New York, 1972.

(3) Beak, P.; Siegel, B. *J. Am. Chem. Soc.* **1976**, *98*, 3601.

(4) Levine, H. L.; Brody, R. S.; Westheimer, F. H. *Biochemistry* **1980**, *19*, 4993.

0002-7863/82/1504-6434$01.25/0