Molecular Simulation of Peptide Interactions with an RP-HPLC Sorbent

I. Yarovsky,† M. T. W. Hearn, and M. I. Aguilar*

Department of Biochemistry & Molecular Biology, Monash University, Wellington Road, Clayton, Victoria 3168, Australia

Received: July 9, 1997; In Final Form: October 2, 1997[⊗]

A simulation procedure based on molecular dynamics has been developed for modeling the interaction of peptides with *n*-alkylsilica reversed phase chromatographic sorbents. A four-step docking procedure was used which included the following stages: (1) interactive rigid-body docking of the peptide with an *n*-butylsilica sorbent using amino acid hydrophobicity coefficients to direct the orientation; (2) automated rigid-body docking by a Monte Carlo simulated annealing procedure in the space of six orientational parameters; (3) solvation of the peptide−sorbent complex with water, and (4) automated docking by molecular dynamics simulated annealing in the full Cartesian coordinate space. The procedure has been validated with the simulation of the binding of the peptide bombesin to an *n*-butylsilica C4 sorbent. The results were analyzed in terms of the change in conformation of both the *n*-butyl ligand chains and the peptide solute following peptide docking. These studies demonstrate that the partial helical character of bombesin was maintained throughout the high-temperature annealing. Overall, this investigation demonstrates the potential of molecular dynamics procedures to aid in the elucidation of peptide interactions with immobilized hydrophobic ligands. Moreover, this study provides a systematic approach to investigate the potential role of hydrophobic effects in peptide−surface interactions.

Introduction

Reversed phase high performance liquid chromatography (RP-HPLC) is now extensively used for the analysis and purification of a wide range of compounds. While the theoretical basis of the RP-HPLC of low molecular weight compounds has been well-established, much less information is available about the precise molecular details of the interaction of peptides and proteins with RP-HPLC sorbents. This situation has arisen because of the strong dependence of the chromatographic retention and bandwidth properties on the three-dimensional structure of these biomacromolecules and the time-dependent changes in conformation that can occur under the commonly used experimental conditions. To characterize the chromatographic process at a molecular level, more precise information is required on the nature of the interaction between the immobilized stationary phase ligands, the solute, and the mobile phase. As part of our studies on the elucidation of peptide retention in RP-HPLC, molecular models for chemically immobilized *n*-butyl, *n*-octyl, and *n*-octadecylsilica phases of different ligand densities for each *n*-alkyl chain length were constructed and subjected to molecular dynamics simulation.[1,2] In these studies, the structural and dynamic properties of these *n*-alkylsilicas as models of commonly used RP-HPLC sorbents were analyzed. The results demonstrated that the simulation reproduced the behavior of the *n*-alkyl ligands in a manner that was consistent with different types of spectroscopic data, and therefore these models represent a good starting point for further study of peptide interactions with *n*-alkylsilica RP-HPLC stationary phases. In the present study, these models of RP-HPLC sorbents were used to develop a four-stage docking procedure in order to simulate the interaction of a small peptide with different chromatographic sorbents. For these investigations, bombesin was selected as a model peptide solute, as the chromatographic properties of bombesin have previously been studied in considerable detail.[3,4] A simulated annealing procedure combined with molecular dynamics (MD) simulation was used to explore the conformational space of bombesin to find minimum energy structures consistent with published experimental data on the conformation of bombesin upon interaction with hydrophobic surfaces. The procedures developed in the present study provide a computational basis for further investigations into the mechanism of peptide separations in RP-HPLC.

Experimental Section

Molecular Model. The model *n*-butylsilica surface was constructed as previously described.[1,2] Briefly, the model was characterized by a ligand density of 2.67 μmol/m^2, a surface hydroxyl group density of 3.90 μmol/m^2, an average surface area per ligand of 62.5 Å^2, the dimensions of the surface in the periodic cell of 28.51 × 28.51 Å^2, and the total number of atoms equal to 1315. Computational results were obtained using software programs from Molecular Simulations Inc. (formerly Biosym Technologies, San Diego, CA), with dynamics calculations performed with Discover and graphical displays created using InsightII. All calculations were performed on a Convex C210 supercomputer.

General Procedure for Modeling Peptide Interactions with an *n*-Alkylsilica Sorbent. The general procedure developed for modeling peptide interactions with the *n*-alkylsilica surface consisted of the following consecutive steps illustrated in Figure 1. Initially, molecular dynamics simulated annealing of the peptide conformation was performed in vacuo using available experimental data on the putative three-dimensional structure of the peptide at hydrophobic surfaces in order to establish starting conformations suitable for subsequent docking to the model *n*-alkylsilica surface. The peptide was initially docked to the model sorbent, using experimental data on the preferred orientation of the peptide with respect to a hydrophobic surface. Rigid-body sampling of the configurational space of the peptide

† Current address: BHP Research, 245 Wellington Rd., Mulgrave, Victoria, 3170, Australia.
* To whom correspondence should be addressed.
⊗ Abstract published in *Advance ACS Abstracts,* November 15, 1997.

S1089-5647(97)02232-3 CCC: $14.00

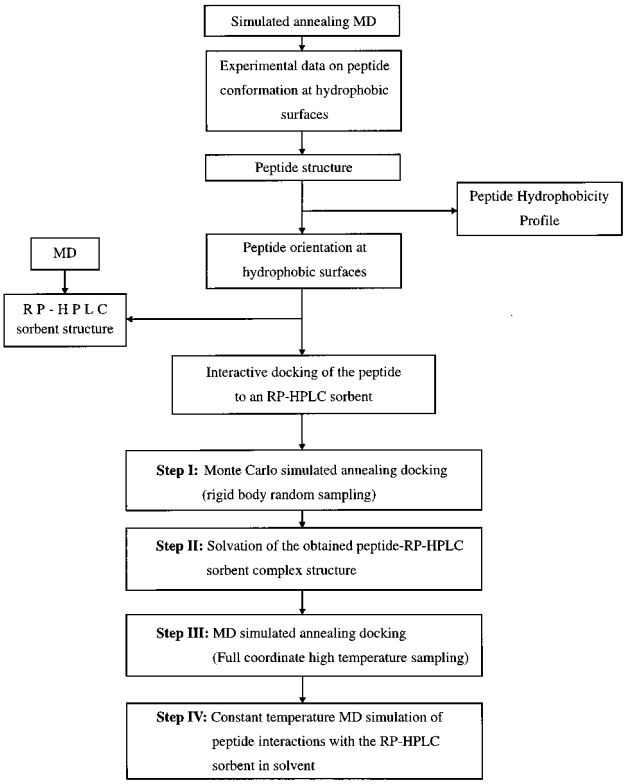

Figure 1. Methodology for the simulation of peptide−sorbent interactions in RP-HPLC.

was then performed by means of a Monte Carlo procedure combined with a simulated annealing algorithm to obtain a starting structure of the peptide−sorbent complex. Subsequent solvation of the model peptide−sorbent complex with explicit solvent molecules was then introduced and the docking procedure continued by means of high-temperature molecular dynamic sampling of the configurational space. Finally, the conformational behavior of the peptide in an RP-HPLC environment was assessed using molecular dynamics at a constant temperature to observe the time development of the system and to collect structural and dynamic data on all system components. This procedure was validated in the present study by the simulation of the interaction of the peptide bombesin with an *n*-butylsilica surface with a ligand density of 2.67 μmol/m^2 in the solvent, water.

(a) Modeling of Bombesin Conformation in the Isolated State. This stage of the simulation was aimed at obtaining an initial conformation of bombesin for further docking to the model *n*-butylsilica surface. The amino acid sequence of bombesin is

Pyr-Gln-Arg-Leu-Gly-Asn-Gln-Trp-Ala-Val-Gly-His-Leu-Met-NH_2

To determine approximate minimum energy conformations that were consistent with the known structure of bombesin in solution and at hydrophobic liquid/solid interfaces, namely the presence of an α-helical structure for the C-terminal residues (Gln6-Met14) and an extended structure for the rest of the peptide,[5] a conformational search was performed by means of a molecular dynamics simulated annealing method with randomly generated structures of bombesin. The potential energy function used in the simulation included bond length stretching and bond angle bending terms, a torsion twisting term, an out-of-plane deformation term for planar systems, and Lennard-Jones and Coulombic terms describing the nonbond van der Waals and electrostatic interactions. The functional form of the energy expression was determined by the CVFF force field[6] employed in the simulation. For the nonbond interactions, a cutoff distance of 12 Å was applied; that is, the nonbond interactions for pairs of atoms separated by distances greater than this cutoff value were neglected. The neighbor list was updated every 20 steps of the calculation. In this initial stage, the simulations were carried out in vacuo with a distance-dependent dielectric constant which accounts for the screening effect of solvent when explicit solvent molecules are not included in the simulation.[7] The peptide bond dihedral angles, ω, were set equal to 180°. To prevent cis−trans isomerism of the peptide bonds at the high temperature, a torsional restraint of 100 kcal mol^{-1} rad^{-2} was added to the peptide bonds.

To achieve conformations for bombesin that satisfied the above criteria, the following simulation protocol was used. Ten randomly generated conformations as well as one linear and one partially helical conformation (comprising residues Gln6-Met14) were used as the initial structures. The potential energies of the initial conformations were first minimized by steepest descent followed by the conjugate gradient minimization method until the maximum derivative decreased to ≤0.5 kcal/(mol Å), thus removing the strains which existed in the randomly generated structures. Molecular dynamics simulated annealing was then performed for the 12 structures. The time step for these molecular dynamics simulations was 1 fs. Initial velocities were generated from a Maxwellian distribution at 1000 K, and the systems were equilibrated at this temperature for 10 ps. Molecular dynamics at constant temperature was then performed at 1000 K for another 10 ps, whereby each system was weakly coupled to a thermal bath using the Berendsen method.[8] The system was then cooled to 300 K by linearly decreasing the temperature by 100 K decrements, allowing the system to equilibrate at each temperature for 10 ps. Starting from the temperature $T = 500$ K, the backbone torsions of amino acid residues 8−14 of bombesin were forced to the helical values (in accordance with experimental observations) using a harmonic force of 20 kcal mol^{-1} rad^{-2}. A conformational search was performed for a total 70 ps (i.e. 10 ps at each temperature), and the conformations with lowest potential energy from the trajectory generated at each temperature were saved on disk for further unrestrained minimization until the maximum derivative was less than 0.01 kcal/(mol Å).

The energies of all the structures generated by the above procedure were then analyzed and the conformations compared using interactive computer graphics methods. The simulated annealing experiment that resulted in a structure that most closely resembled the conformation determined by NMR was then selected. This simulated structure was obtained from the initial partially α-helical conformation. The energy profile of this annealing experiment converged to the minimum energy conformation, which was about 6 kcal/mol less in energy than the conformations that resulted from all other initial structures (Figure 1). Significantly, this local minimum energy conformation correlated with published data on bombesin structure in the presence of hydrophobic surfaces, i.e. the presence of the C-terminal region (residues 8−13) in an α-helical conformation while the N-terminal region was in an extended conformation.[5] However, the obtained conformation cannot be considered as the conformation of the global energy minimum, as an exhaustive conformational search was not performed. The aim of this stage of the development of the peptide−ligand model was to obtain a local energy minimum conformation for bombesin that would be in agreement with independent experimental data on the three-dimensional structure of bombesin upon interaction

with a hydrophobic surface. The molecular dynamics simulated annealing procedure described here allowed this objective to be achieved with relatively small computational expense.

(b) Docking of Bombesin to an *n*-Butyl RP-HPLC Sorbent. The choice of a docking algorithm depends on the availability of experimental information on the mutual orientation of complex components, availability of methods to perform an analysis of electrostatic and/or hydrophobic properties and to correlate them with solvent accessibility of particular atomic groups, and the overall size of the simulated system. In the present study, a combination of different docking algorithms has been used. A four-stage procedure for mutual adaptation of the simulated structures of the immobilized *n*-butyl ligands on the silica surface at an average density of 2.67 μmol/m^2 and the bombesin molecule was developed and consisted of the following steps.

Step I: Interactive rigid-body docking of the peptide with the chromatographic surface, modified with hydrophobic *n*-butyl ligands, carried out in the space of six orientational parameters. Experimental data on preferred peptide orientation with hydrophobic surfaces, the hydrophobic properties of the amino acid residues of bombesin, and interactive intermolecular energy evaluations as described below were used.

Step II: Automated rigid-body docking by means of a Monte Carlo simulated annealing procedure in the space of six orientational parameters using the orientation obtained by step I as the starting structure.

Step III: Solvation of the peptide−sorbent complex.

Step IV: Automated docking by means of a molecular dynamics simulated annealing method in the full space of Cartesian coordinates of both molecules solvated with explicit water molecules.

Step I: Initial Positioning of Bombesin. It is well-established that peptides and proteins interact with reversed phase chromatographic surfaces via specific contact regions in an orientation specific manner.[9−11] The aim of this first step was to determine the initial position of bombesin at the surface of the chromatographic ligands consistent with the published experimental data on the specific orientation with respect to hydrophobic surfaces as well as the hydrophobic properties of bombesin. The initial conformation of the *n*-butyl ligands on the silica surface was taken from the later stage of molecular dynamic simulation performed as described previously.[1] The initial conformation of bombesin was also simulated by molecular dynamics combined with simulated annealing procedures as described above. This simulated conformation of bombesin was used for initial positioning of the molecule in the vicinity of the *n*-butyl ligands on the silica surface. In particular, the C-terminal portion of the peptide which was shown to be in an α-helical conformation was positioned in close contact with the hydrophobic surface as proposed for the orientation of bombesin at the surface of a membrane lipid bilayer, with the Trp-8 residue inserted into the hydrophobic membrane compartment,[5] while the unstructured N-terminal region was exposed to the aqueous solution.

To include the hydrophobic characteristics of bombesin as part of the interactive docking procedure, the amino acid hydrophobicity coefficients derived by Wilce et al.[12,13] were utilized. Amino acid hydrophobicity coefficients determined from reversed phase chromatographic data incorporate information on both the relative solvent transfer characteristics of the amino acid side chains from a polar (water-rich) to a nonpolar (lipophilic) phase as well as information related to the surface accessibility of individual amino acid side chains within a peptide (or protein) structure as it approaches the interactive surface.[12,13] Use of these coefficient scales thus provided a measure of the hydrophobicity and accessible surface areas of the side chains of the individual amino acid residues in specific chemical microenvironments. As a consequence, RP-HPLC-derived hydrophobicity scales of amino acids are potentially extremely useful for interactive docking experiments as they are derived under chromatographic conditions similar to those modeled in the present study. This information allowed the available conformational space for the orientation of bombesin to the *n*-butylsilica surface to be reduced. In particular, a high hydrophobicity value for an amino acid in a specific scale means that, for the specific chromatographic conditions under which this scale was derived, there exists a high affinity of the amino acid to the hydrophobic *n*-alkyl modified surface, which will likely result in the close contact of the amino acid with *n*-alkyl ligands. In contrast, a low hydrophobicity value indicates that in these chromatographic conditions the amino acid is most likely exposed to the aqueous solvent rather than in contact with or buried within the *n*-alkyl phase.

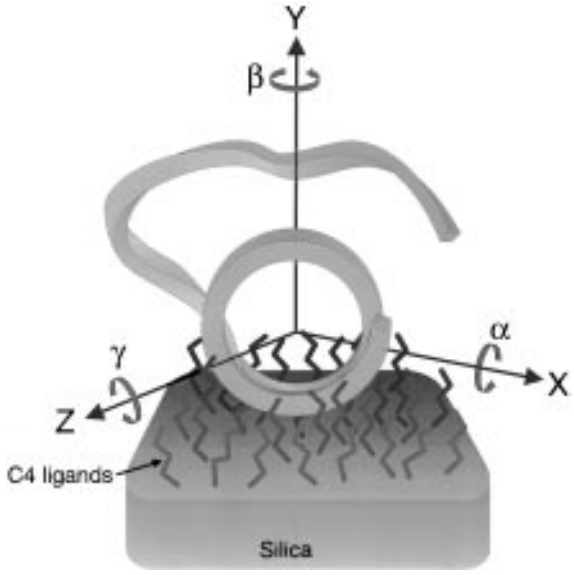

Figure 2. Orientational coordinates for rigid-body docking illustrating the translational coordinates in the *X*, *Y*, and *Z* direction and the rotational coordinates α, β, and γ.

For the interactive docking of the peptide to the *n*-butyl ligands on the silica surface, bombesin was color-coded according to chromatographically derived amino acid affinity coefficients in different *n*-alkyl-modified stationary phases.[13] During the interactive docking procedure, hydrophobic residues, as defined by the RP-HPLC-derived hydrophobicity scale with an *n*-butylsilica sorbent, were positioned closer to the hydrophobic ligands, while hydrophilic residues were oriented away from the *n*-butyl ligands; that is, these residues were more exposed to the solvent.

The next step involved the positioning of the peptide with respect to the surface in a way which minimized the interactive energy but maintained the overall orientation obtained in the previous step. Six independent orientational coordinates are required for full definition of mutual orientation of two molecules in 3D space: three translational (x,y,z) for translations of the center of mass of a guest molecule, and three rotational (α,β,γ) for rotations of the guest molecule about a coordinate axis as depicted in Figure 2. At this stage of the procedure, the intermolecular interaction energy (with the van der Waals and the electrostatic terms applied with a 10 Å cutoff) was gradually minimized interactively by manipulating only the y and β coordinates of the guest molecule (see Figure 2), which

allowed the orientation of the peptide with respect to the surface, based on the above hydrophobicity considerations, to be maintained. The resulting structure was taken as a starting orientation of the bombesin−*n*-butyl ligand complex for the subsequent automated docking procedure.

Step II: Automated Rigid-Body Docking of Bombesin to the RP-HPLC Sorbent. The second (automated) stage of the docking procedure involved the use of a Monte Carlo method for rigid-body adaptation of host−guest molecules in the space of the six independent orientational coordinates (Figure 2). The Monte Carlo method was used for positional searching where the trial molecule performs a random walk in the space around the macromolecule, which is kept static throughout the simulation.[14,15] Using the initial conformation of the complex obtained during step I, the *n*-butylsilica sorbent was immobilized in space and the internal coordinates of both the sorbent and the peptide were held constant. van der Waals and Coulombic energy terms were used for evaluation of the intermolecular interaction energy within a cutoff distance of 10 Å. The Metropolis Monte Carlo method was used[16] for the rigid-body docking of guest−host molecules. A Markov chain of states was generated with one of the independent orientational coordinates (Figure 2) randomly chosen and given a random increment at each step. The intermolecular interaction energy for a new state was calculated, and the probability of the transfer from the old state *i* to a new state (i+1) was determined.[17]

To improve the efficiency of the Monte Carlo search for a minimum energy configuration of the above complex, the standard Metropolis Monte Carlo method was combined with a simulated annealing technique. The procedure started with the high initial temperature of $T = 1000$ K, which allowed more efficient sampling of configuration space. The temperature was then gradually decreased by 50 K every 1000 steps down to $T = 300$ K. The gradual decrease in temperature leads to the progressive narrowing of the configurational space with eventual convergence to the minimum energy area.[14,15,17] The ratio of accepted/rejected configurations was taken as the criterium for convergence (less than 1% of accepted configurations). The procedure was then restarted from another orientation obtained by interactive docking, and after many restarts the configuration with the lowest energy was chosen for further refinement with a full-coordinate docking procedure.

Step III: Peptide−Sorbent Solvation. The complex between the bombesin molecule and the *n*-butyl modified surface obtained by the rigid-body docking procedure was solvated using periodic boundary conditions with unit cell dimensions 28.51 × 28.51 × 38.0 Å. At this stage, 367 water molecules were added to the system, resulting in the total number of atoms in the system equal to 2640. The peptide was surrounded by a water layer of about 7 Å thickness at each side, except for the atoms facing the *n*-butyl chains.

Step IV: Automated Full-Coordinate Docking of Bombesin to the RP-HPLC Sorbent. Any rigid-body docking procedure has an important limitation: the use of static structures of the host and guest molecules. The structures studied here are extremely flexible; thus full internal relaxation must be considered in the procedure of mutual adaptation of the *n*-alkyl ligands and bombesin. All the previous steps of docking thus represent the stage of the model building procedure which creates initial complexes for further investigations. The last step of the docking procedure is the full-coordinate MD relaxation of the system consisting of the *n*-alkyl-modified silica sorbent and the peptide solvated with explicit solvent molecules.

In the following procedure, all the silica surface atoms were kept immobilized, while the surface hydroxyl groups, the *n*-alkyl chains, the peptide, and the solvent atoms were allowed to move. The backbone torsions of bombesin were restrained using a harmonic force of 50 kcal mol^{-1} rad^{-2}. The CVFF force field was used for the potential energy calculations, which included all bond terms, torsion terms, out-of-plane terms as well as nonbond van der Waals and electrostatic terms. The cutoff distance of 12 Å was used for the calculation of the nonbond interaction energy with the neighbor list updated every 20 steps. A dielectric constant of $\epsilon = 1$ was used in the calculation of the electrostatic energy. The integration of the classical equations of motion was performed with a time step of 1 fs during MD.

To improve the mutual adaptation of the peptide and *n*-butyl chains during the automated docking, the following simulated annealing procedure was added to the molecular dynamics procedure. Initial energy minimization by means of steepest descent followed by the conjugate gradient method was used to first relax the system from the initial strain until the maximum derivative decreased to 0.5 kcal/(mol Å). The system was then equilibrated at the low temperature $T = 50$ K for 10 ps and then heated to the temperature $T = 1000$ K by increasing the temperature by 25 K every 1 ps and equilibrating at each temperature. The configurational space of the system was sampled at $T = 1000$ K for 25 ps including equilibration. Following high-temperature dynamics, slow cooling of the system was performed by decreasing the temperature by 25 K every 2 ps and equilibrating at each temperature until the temperature reached 300 K. The energy difference between this simulated structure and that obtained by the rigid-body docking (i.e. from step II) was 36 kcal/mol. Thus, the procedure described above resulted in a significant decrease in the potential energy of the system, which reached the minimum at the lowest sampled temperature. Therefore, after reaching the temperature $T = 300$ K, the docking stage of the procedure was complete.

Inspection of the resulting structure of the system (not shown) revealed that the bombesin retained the C-terminal α-helical structure and the N-terminal extended structure; that is, the conformation of the simulated molecule approximated the secondary structure observed experimentally.[18] In this conformation, the hydrophobic residues of the peptide (according to the C4 hydrophobicity scale[13]) were oriented toward the *n*-butyl chains, while the hydrophilic amino acids which formed the extended structure were exposed to the solvent. It was also apparent that in comparison to the initial structure an improved mutual adaptation of the peptide and hydrophobic chains had been achieved; that is, the C-terminal hydrophobic residues (Trp-8 and His-12) penetrated deeper into the hydrophobic chain layer. This orientation resulted in the creation of a unique arrangement which was consistent with the lipid-induced α-helix found experimentally.[18] This conformation of the peptide−sorbent complex was then submitted to unrestrained constant volume/constant temperature molecular dynamics for data collection and analysis. The *n*-butyl chains, surface silanol groups, the peptide, and 367 water molecules were allowed to move with the only constraint applied to the peptide bond torsion angles, forcing them to the value of 180°. The time development of the system at 300 K was simulated for 200 ps in total, with 100 ps for the equilibration stage and 100 ps for data collection. The equilibrium was checked by monitoring the potential energy and the temperature of the system, which exhibited only small fluctuations after about 80 ps of the equilibration run. During the data collection stage the atomic coordinates and velocities were stored on disk every 100 time steps (0.1 ps) for analysis.

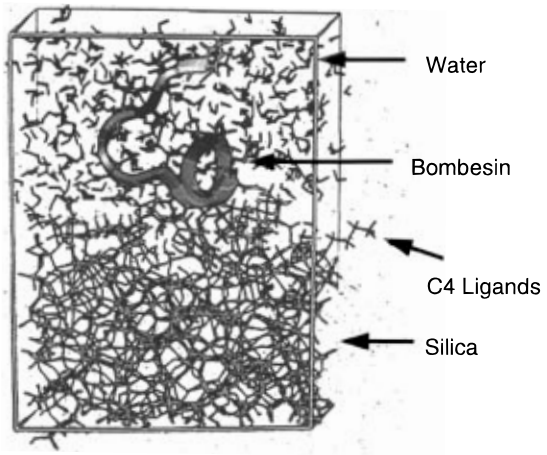

Figure 3. Typical structure of the bombesin−RP-HPLC sorbent−water complex obtained by the constant temperature/constant volume molecular dynamics.

TABLE 1: Fraction of Gauche Conformation for Each Dihedral Angle along an *n*-Butyl Ligand

dihedral number[1]	complex system	in vacuo system
1	0.385	0.615
2	0.462	0.050
3	0.231	0.080
4	0.154	0.230

Results and Discussion

(a) Properties of the *n*-Butylsilica RP-HPLC Sorbent. The structure of the bombesin−*n*-butylsilica sorbent−water complex after unrestrained molecular dynamics is presented in Figure 3. A number of structural and dynamic properties of the *n*-butyl ligands immobilized on the silica surface were derived and included the trans−gauche statistics of chain torsion angles, the thickness of the hydrocarbon layer, the surface accessibility, and the diffusion coefficients for individual carbon atoms. These properties were obtained from the MD trajectory by the methods described in the previous study on *n*-alkyl chain simulations in vacuo.[1,2]

(i) Trans−Gauche Stastistics. The overall proportion of gauche conformations for the *n*-butyl chains in the solvated peptide−*n*-butyl ligand complex was equal to 0.19, which was higher than that found (0.15) for the corresponding in vacuo system.[1] This increase in gauche population resulting from the influence of the surrounding water solvent caused a more compact conformation of the chains presumably due to reinforcement of the intra- and interchain interactions mediated by the hydrophobic effect. Table 1 lists the average fractions of the gauche conformation for each dihedral angle along the *n*-butyl ligand based on the solvated and in vacuo models. As evident from these results, significant differences in the geometry of the ligand attachment to the silica occurred. For example, the proportion of gauche conformations within the complex for torsion number 1 decreased by almost a factor of 2, resulting in the loss of the perpendicular orientation of the ligand with respect to the silica surface. Moreover, the gauche population of torsion number 2 increased dramatically, also leading to the high chain tilt. Such a distribution of gauche defects along the *n*-butyl chains demonstrates that the most "broken" part of a chain is now near the silica surface, which leads to the highly tilted orientation of the chains with respect to the silica surface. The predominance of the end-gauche conformation which was established in the in vacuo system[1] disappeared in the presence of the polar solvent. Instead, the gauche conformation was more uniformly distributed along the *n*-butyl chain, allowing a more compact chain conformation. This result is consistent with the consequences of the hydrophobic effect (that is, the *n*-butyl chains have minimized their exposure to the water molecules) and is in agreement with the theoretical prediction by Pratt and Chandler[19] of an increase in the gauche population upon transfer of *n*-butane from the gas phase to aqueous solution.

(ii) Hydrocarbon Layer Thickness and Surface Accessibility. The structure of *n*-butyl ligands in the solvated complex discussed in the previous section in terms of trans−gauche statistics was further confirmed from the calculated hydrocarbon layer thickness. This parameter decreased significantly in comparison with the in vacuo simulation, resulting in a value of 3.85 Å compared to 5.45 Å in vacuo.[1] This result shows that the chain conformation was more bent in the presence of the solvent and the peptide, with the *n*-butyl chains positioned closer to the silica surface. This arrangement of the ligands in the compact gauche conformation allows them to be accommodated almost parallel to the surface at the given surface density when the near neighbor spacing is equal to 7.9 Å. The highly tilted or surface parallel chain position results in a better coverage of the silica surface, i.e. in a decrease in the accessibility of the silica surface, including unreacted surface silanols, to the solvent. The solvent accessible surface area was calculated using the Connolly algorithm with a water probe (as for the in vacuo systems[1]). It was found that when covered by the ligands the solvent accessible surface area decreased by 9.57% (for comparison, at the same ligand density for the in vacuo system, it decreased by 6.63%), demonstrating the changed chain position with respect to the surface and improved surface protection in the peptide−ligand complex.

(iii) Distribution of the Solvent around n-Butyl Ligands and Hydrogen Bonding to the Unreacted Silanols. Analysis of the hydration of *n*-butyl ligands can provide important information on the contribution of the hydrophobic effect to the molecular mechanism of RP-HPLC. Although conformational equilibria of hydrocarbon chains in water and in other solvents have been investigated by many simulations,[20] the results have not been conclusive. It is generally acknowledged that the molecular orientation of liquid water on hydrophobic surfaces as well as around small hydrophobic solutes is governed by the same principle, namely, the optimization of hydrogen-bonding interactions of water which results in the so-called hydrophobic effect. However, the molecular mechanism of the hydrophobic effect for small nonpolar solutes is different from the solvation of larger nonpolar surfaces. In particular, the hydration of small nonpolar solutes is characterized by a clathrate-like geometry for the solvent, driven by the large enthalpic penalties that would follow from a less ordered solvation layer. In such structures, the orientation of water molecules near the solute is biased so that hydrogen-bonding groups (donor or acceptor) tend to avoid interaction with the nonpolar solute, which cannot itself participate in hydrogen bonding. The solvent thus maintains a hydrogen-bonding interaction that is comparable to that in bulk solution. In a study of the structure of liquid water at an extended hydrophobic surface,[21] it was found that the liquid structure nearest the surface is characterized by "dangling" hydrogen bonds; that is, a typical water molecule at the surface has one potentially hydrogen-bonding group oriented toward the hydrophobic surface. This surface arrangement represents a balance between the tendencies of the liquid to maximize the number of hydrogen bonds, on one hand, and to maximize the packing density of the molecules on the other. Thus, the hydration structure of large hydrophobic surfaces can be very different from that of small hydrophobic molecules. The *n*-alkyl

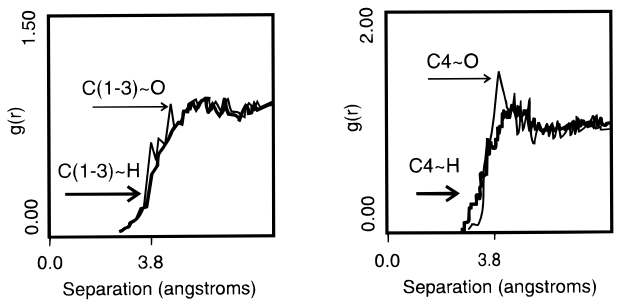

Figure 4. Radial distribution functions for *n*-butyl ligands and water. C(1−3) = the methylene (CH_2) groups and C4 = the terminal methyl (CH_3) groups of the *n*-butyl ligands; O = oxygen; and H = hydrogen of water molecules.

ligands immobilized on a silica surface cannot be considered either as small hydrated hydrophobic solutes or as an extended hydrophobic surface due to the incomplete coverage of the polar silanols by the *n*-butyl ligands. Thus, the mechanism of their hydration in the presence of amphoteric solutes such as peptides may be different from those described above.

To gain insight into the hydration mechanism of *n*-butyl ligands that form part of a solvated peptide−sorbent complex, ligand−solvent radial distribution functions (RDF) for methyl and methylene groups with the oxygen and hydrogen atoms of water were constructed. The results are shown in Figure 4, with the data corresponding to the ligand methylene groups (CH_2 labeled C(1−3) in Figure 4)−water distributions provided in the left-hand panel and terminal methyl groups of the ligand (CH_3 labeled (C4) in Figure 4)−water distributions given in the right-hand panel. It can be seen that the CH_3−O RDF has the most structure, with the first maximum at 3.7 Å. This value is similar to the RDF reported by Jorgensen for *n*-butane in water.[22] In contrast, the CH_2 groups show little structure in their RDFs (Figure 4), which demonstrates that they are less exposed to the solvent than CH_3 groups. The first peak in the CH_2−O RDF occurs at 3.95 Å and reveals that there are closer contacts between water molecules and the CH_3 groups than with the CH_2 groups. From the CH_3−H RDF, it is evident that the positions of the hydrogen atoms of water around the alkyl groups are more random than for the oxygen atoms. It is also apparent that some hydrogen atoms from the first coordination shell are closer to the alkyl groups than the oxygen atoms. A similar situation was also found for *n*-butane in water.[22]

To assess the degree of protection of the surface silanol groups from the interactions with the solvent, the formation of hydrogen bonds between water and surface hydroxyl groups was also examined. From the complex structure obtained by MD (Figure 3), it was found that about 16% of surface silanols are involved in hydrogen bonding to water. These data show that surface silanols are reasonably protected by the *n*-butyl ligands.

In summary, it can be concluded that the molecular picture of water around the *n*-butyl ligands in an *n*-butylsilica environment is closer to that obtained for *n*-butane in water, rather than for water structure on an extended hydrophobic surface. This result is reasonable because the *n*-butyl ligands in such RP-HPLC environments are immobilized on the silica surface at a relatively low density (i.e. 1−3 μmol/m^2) and thus do not generate a homogeneous hydrophobic surface. However, an increased ordering of the water molecules in the first shell around the *n*-butyl ligands was observed which optimizes the hydrogen-bonding interactions, thus consistent with the participation of the hydrophobic effect.

(iv) Mobility of the Ligands. The mobility of the *n*-butyl ligands in the solvated peptide−ligand complex was studied in terms of the dynamic behavior of the individual carbon atoms as described for the in vacuo system.[1] Diffusion coefficients were calculated from the average mean squared displacement plots using the Einstein relation. Table 2 lists the diffusion coefficients of individual carbon atoms of *n*-butyl ligands in the complex system as well as in vacuo for comparison. It can be seen that in the solvated peptide−sorbent complex the mobility increases from the C1 to C4 atom in a similar fashion to that observed for all three systems in vacuo.[1] However, the increase was much greater than observed previously for the in vacuo system; in the solvated peptide−ligand complex, the mobility of the C4 atom was larger by almost a factor of 10 than that of the C1 atom. The overall mobility of the C1 to C4 atoms in the ligand−solute−solvent complex is higher than in vacuo for all the carbons except for the attachment carbon at the surface (C1), where the mobility is almost the same in both systems. The increase in mobility in comparison to the in vacuo system is especially pronounced for the terminal C3 and C4 atoms; that is, for the C4 atom, the difference was a factor of 2.2, for the C3 by a factor of 3.9, and for C2 by a factor of 1.5. Such an increase in the mobility can be explained by the influence of the solvent and the peptide which cause the *n*-butyl chains to rearrange, leading to the increased proportion of gauche defects. The unchanged mobility of the C1 atom in the solvated peptide−ligand complex compared to the *n*-butylsilica in vacuo may be explained by the significant restraints imposed on the C1 atom by the covalent bond to the surface as well as the fact that solvent molecules are not able to penetrate sufficiently deeply into the hydrophobic chain layer.

TABLE 2: Diffusion Coefficients of Individual Carbon Atoms of *n*-Butyl Ligands (10^{-6} cm^2/s)

carbon atom number[1]	complex system	in vacuo system
C1	0.17	0.18
C2	0.35	0.22
C3	1.10	0.28
C4	1.47	0.67

In summary, the mobility of the *n*-butyl chains is higher in the sorbent−solvent−solute system than in vacuo, with the most pronounced increase for the terminal atoms. The overall mobility of individual carbons increased with the distance from the silica surface, as was also established for the corresponding *n*-butylsilica in vacuo.[1]

(b) Properties of the Solute. *(i) Conformation of Bombesin upon Interaction with the RP-HPLC Sorbent.* The conformation of bombesin obtained after the molecular dynamics and simulated annealing procedure in the RP-HPLC environment is presented in Figure 3. Upon interaction with the *n*-butyl silica sorbent, bombesin maintained a partially helical comformation, with an α-helical segment at the C-terminus of the peptide and a random coil conformation at the N-terminus. The following hydrogen bonds between bombesin residues could be identified: Leu4−Gln7 and Asn6−His12 for the side chain donor−acceptor pairs, as well as Leu4−Asn6, Gln7−Gly11, Gln7−Val10 for backbone N---O pairs. The hydrogen bonding between the Asn6 and His12 side chains may influence the stabilization of the secondary structure in the C-terminal region of the peptide and play a role in the biological activity of the peptide. The hydrogen bonding between Trp8 and His12 suggested by Erne and Schwyzer[18] was not evident. However, the stabilizing effect of this H-bond on the membrane structure of bombesin may be substituted by the Asn6−His12 hydrogen bond established during the MD simulation. It is interesting to note that the absence of a hydrogen bond between the side chains of Trp8 and His12 was also established by ^{1}H NMR spectroscopy experiments on bombesin in trifluoroethanol/water mix-

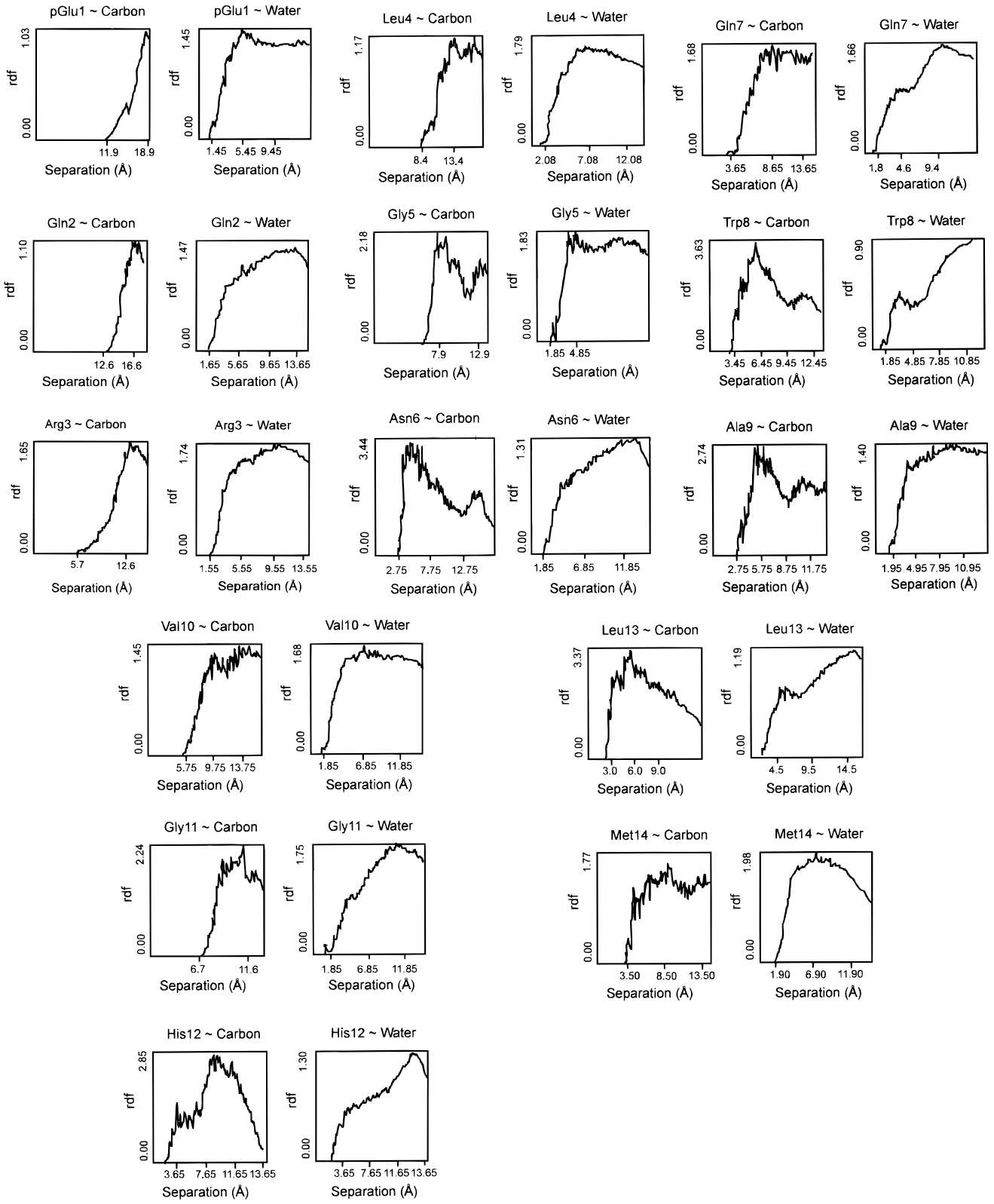

Figure 5. Radial distribution functions of hydrocarbon chain (left-hand-side figures) and water atoms (right-hand-side figures) around each amino acid residue in bombesin.

tures.[23] The conformation obtained in the present study allows the peptide to be oriented by its hydrophobic residues toward the hydrophobic surface with the Trp8 and His12 residues being inserted in the hydrocarbon compartment created by the *n*-butyl chains. This orientation is consistent with the behavior observed experimentally for bombesin upon contact with the membrane lipids.[18] It should be noted that the ranking of histidine and tryptophan as highly hydrophobic in the C4 scale in contrast to the hydrophilic ranking with other RP-HPLC scales has been previously discussed[13] and relates to the ability of the amino acids to interact with both hydrophobic ligands and the surface silanols, thereby enhancing their overall interactive potential with *n*-butylsilica sorbents.

(ii) Orientation of Bombesin with Respect to the Surface. The docking problem in molecular modeling arose about 15 years ago when the first attempts were made to explain the structure−

activity relationship of a drug in terms of molecular structure using real-time interactive computer graphics.[24−29] Today docking is widely used in computer-aided drug design as the basis of the biological activity of most drug molecules as found in their mode of interaction with a specific receptor molecule. In these applications with protein−ligand interactions, steric fit and electrostatic interactions are the two main criteria for maximizing drug−receptor interactions. However, hydrophobic interactions also play an important role in a wide range of peptide and protein surface interactions, but cannot be readily quantitated, and as a consequence the energetic contributions of hydrophobic interactions to a particular molecular interaction cannot be readily assessed. As the molecular basis of hydrophobic interactions is poorly understood in terms of a precise physical model to describe the free energy of these interactions, the present study provides a potentially useful approach to simulate the hydrophobic interactions associated with the RP-HPLC of peptides.

Hydrophobic characteristics of molecules have previously been used in interactive docking experiments[30,31] in assessing their likely binding sites and modes of interactions. In 1986, the concept of a hydrophobic dipole moment was introduced[32] as a measure of the asymmetry of molecular hydrophobicity. This approach was extended by Furet and co-workers and Brasseur, who proposed a three-dimensional representation of the hydrophobicity profile that allows a detailed description of the molecular hydrophobic characteristics in terms of hydrophobic and hydrophilic regions on the van der Waals surface.[33,34] As utilized in the present investigation, the visualization of the hydrophobic and hydrophilic envelopes around the molecular surface enables this property to be used in interactive docking procedures. This method is very useful for finding the initial orientation of a guest molecule before applying molecular dynamics calculations since the available space to be explored is significantly and rationally narrowed, thus reducing the computational cost of experiments.

In automated docking[35−38] the interacting molecules are positioned according to algorithms that vary from exhaustive search[39−41] and distance geometry methods[42−44] to stochastic approaches.[14,15] Compared to manual docking, automated methods are less dependent upon the user's considerations about the most important areas of binding as well as allowing the exploration of a large area of conformational space. Nevertheless, a refinement of the complex structures so obtained is also necessary.[15,45]

To study the mechanism of molecular interactions of bombesin with the RP-HPLC sorbent and solvent, the distribution of carbon atoms of the *n*-butyl ligands as well as the structure of water molecules around different amino acid residues of bombesin was investigated. In particular, radial distribution functions of water oxygens and hydrogens as well as *n*-butyl chain carbon atoms around specific peptide side chains were computed and are presented in Figure 5. These data allowed the overall exposure of the peptide amino acid residues to water and/or the hydrophobic chains to be characterized.

In addition, the relative solvent accessibility for each residue of bombesin in the conformation adopted upon binding to the C4 sorbent was calculated as a ratio of solvent accessible area determined using the Connolly algorithm[46] and the theoretical maximum accessible surface area determined by Rose.[47] A low accessibility indicates that an amino acid residue is buried within the secondary structure of the peptide, while a high accessibility indicates that the residue can be in a close contact with a surrounding environment, whether it is the solvent or the stationary phase ligands. The data for the relative residue accessibility (Figure 6) were used together with RDFs in the detailed analysis of bombesin residue interactions with different phases of the chromatographic environment.

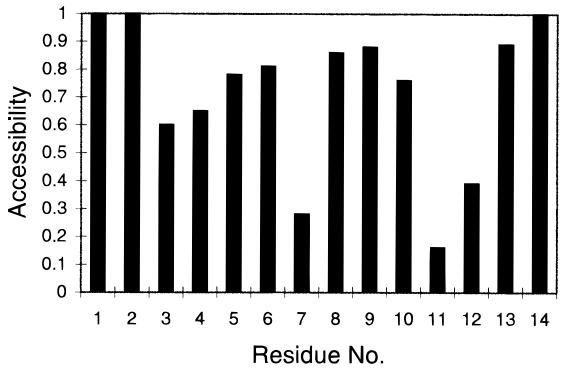

Figure 6. Relative surface accessibility of bombesin residues in a typical conformation adopted upon binding to the RP-HPLC sorbent.

Detailed analysis of the radial distribution functions for carbon atoms and water molecules around each residue of bombesin (Figure 5) showed that residues 1−5 were highly exposed to the solvent as the first maximum in residue−water RDFs occurs at about 2 Å. This result may arise from the presence of water around the two polar residues Gln2 and Arg3, which may be involved in hydrogen bonding with water molecules. In contrast, the residue/carbon RDF exhibits the first maximum at much longer distances of 8−15 Å, which demonstrates that there were no ligands in contact with residues 1−5. For the Asn6 residue, there is a broad peak for carbon at 2.75−6.0 Å together with a broad distribution of water, which shows that the residue was in contact with both phases. This behavior is possibly due to the existence of both the methylene side chain and polar amide group in this residue, which is also highly accessible for interactions with the environment (Figure 6). The next residue, Gln7, was also found to be exposed to the ligands and to water but to a lesser extent than Asn6, consistent with its lower overall accessibility (Figure 6). The Trp8 residue exhibited a sharp, high peak in its carbon RDF, which demonstrates that there was a structuring of carbon chains around this residue. Together with a high accessibility of this residue (87%) shown in Figure 6 these results are consistent with the partitioning of the residue into the hydrocarbon phase. The significance of this observation on the molecular mechanism of the bombesin−*n*-butyl ligand interactions can be emphasized since it reflects the important role of Trp8, already established to be essential for the biological activity of this peptide. There was almost no water predicted by the RDF analysis to be around this residue (Figure 5). This figure also reflects the hydrophobic association between the residue and the *n*-butyl ligands. The next residue in the bombesin sequence, Ala9, also demonstrated the existence of hydrocarbon structure around it, exhibiting a broad distribution in the carbon−Ala9 RDF at 2.75−8.0 Å. Water is only present at distances from about 4 Å. Because Ala9 is hydrophobic and highly accessible to the *n*-butyl ligands (Figure 6), this residue is thus able to "support" the strong association of the neighboring Trp8 with the hydrocarbon phase. Val10 and Gly11 were found to be less exposed to the hydrocarbon phase than Trp8 and Ala9 and showed the presence of some surrounding water. However, His12 was again involved in a strong association with the *n*-butyl chains. The first peak in the carbon RDF occurs at 3.65 Å, which demonstrates the close contact of the residue atoms with hydrocarbon ligands. Therefore, despite its relatively low accessibility (Figure 6) His12 may play an important role in the molecular mechanism of the interaction of bombesin with the *n*-butyl ligands and the surface silanols of the chromato-

graphic sorbent presumably through silanol−imidazole interactions. Water was not closely structured around this residue due to its relatively deep penetration into the hydrocarbon layer. There was a high degree of structuring of the hydrocarbon chains around the next highly hydrophobic and highly accessible (Figure 6) residue, Leu13, corresponding to the high and broad maximum in carbon−Leu13 RDF at 3.5−8.0 Å, as shown in Figure 5. Water−Leu13 RDF exhibited a low first peak at about 4.5 Å, demonstrating no significant interaction of water with this residue. The C-terminal Met14 residue is 100% accessible and highly hydrophobic. It showed a broad carbon distribution starting from 3.5 Å as well as a broad water distribution from about the same separation distance (about 3.5 Å). The attraction of water may be explained by the neighboring C-terminus and a more disordered side chain structure. However, the residue is also able to contribute to the hydrophobic association with ligands due to its close contacts with carbon exhibited by RDF.

Overall, it should be noted that C-terminal residues of the peptide formed two clusters of hydrophobic residues: Trp8-Ala9-Val10-Gly11 and Leu13-Met14, which ensure a high degree of hydrophobic interaction with immobilized ligands. Those clusters may also "promote" the interaction of the polar His12 with the hydrocarbon phase. Thus, it may be concluded that the individual residues have a different mechanism of molecular interaction with the stationary phase. Some amino acid residues were able to partially penetrate into the nonpolar phase, while other amino acid residues are simply oriented and adsorbed onto the modified silica surface. This mechanism is facilitated by the conformational flexibility of both the peptide and the *n*-butyl ligands. Both processes allow the peptide to adopt a structure that ensures that the strongest interactions occur. These processes are achieved through an increased gauche population of the ligands interacting with water, as well as increased chain atom mobility, which both provide the possibility of hydrophobic association of the hydrophobic peptide residues with the ligands. In the case of the shorter *n*-butyl ligands, these hydrophobic interactions can be further enhanced by polar interactions between amino acid side chains and the surface silanols. These results therefore highlight the molecular complexities of the retention process, which cannot be readily described solely as a solvophobic or a partition process, a conclusion also reached in a recently reported molecular dynamics simulation of the interaction of methane with C18 ligands.[48]

Conclusion

In this study, a general model and a computational procedure have been developed that can be used to simulate the behavior of all the molecular components of an RP-HPLC process at the molecular level. In particular, this study represents the first example in which the docking of a peptide to an RP-HPLC sorbent has been performed using a fully explicit atomistic model of the *n*-alkyl surface and incorporating the hydrophobic and energetic characteristics of the systems. Further time development of the constructed system was simulated by means of a molecular dynamics method. The procedure was validated using a solvated peptide−ligand system which consisted of an *n*-butyl-modified silica sorbent, the tetradecapeptide bombesin, and water as a solvent, and the results of the simulation reproduce the available experimental data on bombesin behavior at hydrophobic surfaces. Thus, the model and procedures developed in this study provide the basis for further characterization of the molecular basis of peptides and proteins in RP-HPLC environments. The results of the present study also provide important insight into the underlying principles of peptide folding and peptide−surface interactions at the atomic level.

Acknowledgment. The support of the Australian Research Council is gratefully acknowledged. The molecular dynamics simulations were carried out on a CONVEX C210 computer located at the Biomolecular Research Institute (Parkville, Australia).

References and Notes

(1) Yarovsky, I.; Aguilar, M. I.; Hearn, M. T. W. *J. Chromatogr*. **1994**, *660*, 75.
(2) Yarovsky, I.; Aguilar, M. I.; Hearn, M. T. W. *Anal. Chem.*. **1995**, *67*, 2145.
(3) Purcell, A. W.; Aguilar, M. I.; Hearn, M. T. W. *J. Chromatogr*. **1992**, *593*, 103.
(4) Purcell, A. W.; Aguilar, M. I.; Hearn, M. T. W. *Anal. Chem.* **1993**, *65*, 3038.
(5) Cavatorta, P.; Farruggia, G.; Masotti, L.; Sartor, G.; Szabo, A. G. *Biochem. Biophys. Res. Commun.* **1986**, *141*, 99.
(6) Dauber-Osguthorpe, P.; Roberts, V. A.; Osguthorpe, D. J.; Wolff, J.; Genest, M.; Hagler, A. T.; *Proteins* **1988**, *4*, 31.
(7) Fritsch, V.; Westhof, E. in *Modelling of Molecular Structures and Properties*; Rivail, J. L., Ed.; Elsevier: Amsterdam, 1990.
(8) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F., DiNola, A.; Haak, J. R. *J. Chem. Phys*. **1984**, *81*, 3684.
(9) Purcell, A. W.; Aguilar, M. I.; Hearn, M. T. W. *Peptide Res*. **1995**, *8*, 160.
(10) Regnier, F. E. *Science* **1987**, *238*, 319.
(11) Zhou, N. E.; Mant, C. T.; Hodges, R. S. *Peptide Res*. **1990**, *3*, 8.
(12) Wilce, M. C. J.; Aguilar, M. I.; Hearn, M. T. W. *J. Chromatogr*. **1991**, *536*, 165.
(13) Wilce, M. C. J.; Aguilar, M. I.; Hearn, M. T. W. *Anal. Chem.* **1995**, *67*, 1210.
(14) Goodsell, D. S.; Olson, A. J. *Proteins: Struct., Funct., Genet*. **1990**, *8*, 195.
(15) Cherfills, J.; Duquerroy, S.; Janin, J. *Proteins: Struct., Funct., Genet*. **1991**, *11*, 271.
(16) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. J. *Chem Phys*. **1953**, *21*, 1087.
(17) Shamovsky, I. L.; Yarovsky, I.; Khrapova, N. G.; Burlakova, E. B. *J. Mol. Struct. (THEOCHEM)* **1992**, *253*, 149.
(18) Erne, D.; Schwyzer, R. *Biochemistry* **1987**, *26*, 6316.
(19) Pratt L. R.; Chandler, D. *J. Chem. Phys*. **1977**, *67*, 3683.
(20) Ohmine, I.; Tanaka, H. *Chem. Rev*. **1993**, *93*, 2545.
(21) Lee, C. Y.; McCammon, J. A.; Rossky, P. J. *J. Chem. Phys*. **1984**, *80*, 4448.
(22) Jorgensen, W. J. *J. Chem. Phys*. **1982**, *77*, 5757.
(23) Carver J. A.; Collins, J. G. *Eur. J. Biochem*. **1990**, *187*, 645.
(24) Porter, T. *Comput. Graphics* **1978**, *12*, 282.
(25) Feldman, R. J.; Bing, D. H.; Furie, B. C.; Furie, D. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 5409.
(26) Smith G. M.; Gund, P. *J. Chem. Inf. Comput. Sci*. **1978**, *18*, 207.
(27) Greer, J.; Bush, B. L. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 303.
(28) Max, N. L. *Comput. Graphics* **1979**, *13*, 165.
(29) Langridge, R.; Ferrin, T. E.; Kuntz, I. D.; Connolly, M. L.; *Science* **1981**, *211*, 661.
(30) Saunders, J.; Findlay, J. B. C. *Biochem. Soc. Symp*. **1991**, *57*, 81.
(31) Aqvist, J.; Tapia, O. *J. Mol. Graphics* **1992**, *10*, 120.
(32) Eisenberg, D.; McLaughlan, A. D. *Nature* **1986**, *319*, 199.
(33) Furet, P.; Sele, A.; Cohen, N. C. *J. Mol. Graphics* **1988**, *6*, 182.
(34) Brasseur, R. *J. Biol. Chem*. **1991**, *266*, 1612.
(35) Connolly M. L. *Biopolymers* **1986**, *25*, 1229.
(36) Goodford, P. J. *J. Med. Chem*. **1985**, *28*, 849.
(37) Lipkowitz, K. B.; Zegarra, R. *J. Mol. Biol.* **1978**, *124*, 323.
(38) Jian, K.; Kim, S.-H. *J. Mol. Biol*. **1991**, *219*, 79.
(39) Wodak, S. J.; Janin, J. *J. Mol. Biol.* **1978**, *124*, 323.
(40) Santavy, M.; Kypr, J. *J. Mol. Graphics* **1984**, *2*, 47.
(41) Goodsell, D.; Dickerson, R. E. *J. Med. Chem*. **1986**, *29*, 727.
(42) Kuntz, I. D.; Blaney, J. M.; Oatley, S. J.; Langridge, R.; Ferrin, T. E. *J. Mol. Biol*. **1982**, *161*, 269.
(43) Billeter, M.; Havel, T. V., Kuntz, I. D. *Biopolymers* **1987**, *26*, 777.
(44) Meng, E. C.; Shoichet, B. K.; Kuntz, I. D. *J. Comput. Chem*. **1992**, *13*, 505.
(45) Jähnig, F.; Edholm, O. *J. Mol. Biol.* **1992**, *226*, 837.
(46) Connolly, M. L. *J. Appl. Crystallogr*. **1983**, *16*, 548.
(47) Rose, G. D.; Geselowitz, A. R.; Lesser, G. J.; Zehfus, M. H. *Science* **1985**, *229*, 834.
(48) Klatte, S. J.; Beck, T. L. *J. Phys. Chem.* **1996**, *100*, 5931.