

# Computational Analysis of Adhesion of Primer Ligands to Dentinal Collagen: Effect of Spacer Groups in Ligand and Amino Acid Residue Sequence Differences in Collagen\*

J. Vaidyanathan,<sup>1</sup> S. Ravichandran<sup>2</sup> and T.K. Vaidyanathan<sup>1</sup>

<sup>1</sup>Department of Restorative Dentistry, University of Medicine and Dentistry of NJ, Newark, NJ 07103; <sup>2</sup>Advanced Biomedical Computing Center, National Cancer Institute, SAIC/Frederick, Frederick MD 21702

**Abstract:** This study sought to assess by computer modeling the interactions between dentinal collagen and primer monomer ligands used to promote bonding of restorations to tooth. Modeling was carried out both by direct and indirect methods to probe interaction mechanisms. Ligands studied in this investigation conformed chemically to methacrylate phosphates of alkane diol, with changes in the number of methylene spacer groups. Increase in number of methylene groups in the series introduces increasing levels of ligand conformational freedom. An automatic docking program was used to analyze the effect of these changes on primer-collagen interactions in direct (target-based) modeling. The effect of limited modifications of amino acid residue sequences in structural variants of type 1 dentinal collagen was also assessed in this approach. The indirect (ligand-based) modeling used a pharmacophore search to mimic primer binding to type 1 collagen using a common functional alignment algorithm. Docking energy, and the non-bonded and electrostatic contributions to it, showed statistically highly significant differences ( $p < 0.0001$ ) with ligand conformational freedom. But the effect of collagen composition differences was, although statistically significant ( $p < 0.05$ ), relatively small. Both target-based direct docking and ligand-based indirect modeling visualizations showed that conformations tended to align in a 3-D geometric pattern in bound states, and that the conformational flexibility of the ligands played a critical role in alignment. The results suggest that incorporation of spacer groups in primer monomers may modify dentin bonding to improve overall adhesion under optimum conditions.

**Key Words:** Adhesion, conformational changes, collagen, interaction energy, dentin bonding, pharmacophore

## INTRODUCTION

Adhesion is of fundamental importance in biological processes involving bonding of adhesives to tissues, drug-protein interaction, bacterial adhesion to host tissues, cell signaling, enzyme catalysis etc. Biological adhesion is often visualized as binding events involving specific target sites with complementary recognition sites of ligand as a lock-and-key interaction event with or without an induced fit [1, 2]. Typically, it is assumed that the interaction event is overwhelmingly favored at a preferred target site, such as a cavity in the target molecular structure, where steric and electrostatic complementarity to a recognition site in the ligand facilitate ligand-target binding leading to the formation of an intermolecular complex favored by energy optimization.

Computer modeling simulations have become valuable tools to study binding in biological processes involving adhesion, adsorption etc. [3]. Such simulations can be direct or indirect. In the direct or target-based simulation, ligand interaction with a target molecule is modeled and the energy of interaction is computed. Such modeling is carried out using

well known molecular modeling software programs such as Dock, AutoDock, Sybyl, Amber and many others programs, as described previously [3]. In the indirect ligand-based method, the common underlying 3-D alignment pattern of the functional features of known bioactive ligands (known as pharmacophore) is generated from the low energy conformations of ligands [4, 5]. In essence, molecules are considered active in a particular target if they carry a number of common features with favorable geometric placements in space to interact favorably with complementary features in a target. This suggests that all or most of the functional groups of the ligand should be aligned in space to their complementary sites on the target molecule for efficient ligand-target interaction. Thus the predominant common functional alignment (or alignments) displayed by low energy conformations of a group of known molecules active to a given target can be thought of as a molecular framework (or frameworks) encoding the binding features necessary to bind to the target under consideration. The indirect method uses bioactive ligands to identify such a molecular framework or template. Both direct and indirect methods are valuable to probe adhesion to biological molecules, and this study used these methods to explore interaction of hydrophilic monomers (used as primers in dentin bonding) with a type 1 collagen target molecule.

A practical application of computer modeling simulations is the case of bonding of restorations to tooth. It is to be pointed out that the term 'bonding' is used in adhesive dentistry to represent a stable attachment, but does not necessarily suggest a primary bond as is traditionally implied chemi-

\*Address correspondence to this author at the Department of Restorative Dentistry, University of Medicine and Dentistry of NJ, Newark, NJ 07103; Tel: -----; Fax: -----; E-mail: jvaidyan@umdnj.edu

<sup>1</sup>This study was in part presented at the International Conference on Computational and Experimental sciences and engineering at Miami Beach, FL in Feb 2007.

cally. For many years, such bonding of restorations to tooth was a serious problem at the restoration-dentin interface. This is because dentin contains an organic matrix of type 1 collagen scaffold and a mineral phase of hydroxyapatite (HAP). The hydrophobic monomers previously used in bonding restorations to tooth failed to efficiently bond to dentin because of (a) the moist environment at the dentin surface due to the dentinal fluid emanating through the dentin tubules and (b) the low surface energy of the organic collagen matrix. Monomers with hydrophilic functionalities have been used in recent years as primers to improve bonding in the hydrophilic environment of native collagen. The overall "bonding" procedure typically involves a combination of etching, priming and bonding steps. In the etching step, the mineral phase of HAP is separated from the collagen fibrils. The priming step involves treatment with a hydrophilic monomer (called primer monomer), which permeates into the dentin sub-surface layer and primes or interacts with the collagen matrix. The bonding step involves use of a hydrophobic monomer (often termed as adhesive monomer) between a restoration and the tooth for simultaneous copolymerization of the primer monomer with the adhesive monomer, and the adhesive monomer with the composite restoration. It has been shown by scanning electron and transmission electron microscopy that the interphase region between the tooth tissue and the restoration contains a collagen-primer resin hybrid layer after bonding [6]. The formation of the hybrid layer at the dentin-restoration interface has been recognized as the seminal event that promotes long-term durability and stability of restoration bonding to tooth [6]. The interaction of primer monomer with collagen during the initial priming step is therefore critically important. The collagen-primer interaction mechanism is not fully understood, and computer modeling is not only a practical tool to understand potential interaction mechanisms, but may also be useful to develop criteria for the design of hydrophilic monomers to be used as primers to improve dentin-bonding. It was shown in our previous reports that ligand binding to collagen molecules may play an important role in dentin bonding [3, 7, 8]. The collagen-ligand intermolecular interaction energy in dentin bonding was computationally estimated using docking simulations of a flexible ligand on a rigid type 1 collagen molecular target [3]. A docking algorithm, which facilitates modeling in a hydration environment of native collagen in tissues with an implicit solvent (water) approach, was used for the analysis. Such simulations have shown that ligand docking is favored to occur as a 3-D alignment of functional features complementary to the configuration of target sites along cavity tracks on type 1 collagen molecular surfaces [3, 7, 8]. The triple helix motif in collagen structure is characterized by cavity tracks on the molecular surface due to the helical turns of the coil-coil structure of collagen molecule. The triple helix motif presents multiple sets of discrete binding sites in these tracks, and most of these binding sites are typically similar due to the repeating amino acid sequences in the structure. However some differences in the composition and topography at the cavity tracks are expected. Thus, the amino acid residues in the collagen molecule follow the repeating order -GLY-X-Y-sequence, where GLY is glycine, X and Y are often proline (PRO) and hydroxyproline (HYP), respectively. However alanine (ALA), glutamic acid (GLU), lysine (LYS) and other

amino acid residues have been shown to be present in dentinal collagen [9]. Adhesion to such a structure is therefore a set of binding events, which mostly occur at similar target sites in the cavity tracks along the length of the collagen molecule. But substitutions to one or more residues in the triplet sequences also lead to a few dissimilar binding sites where binding may be favored. The presence of multiple binding pockets with spatial and charge perturbations due to differences in amino acid residues adds an interesting dimension to the identification and analysis of the binding events in the collagen molecule. Molecular modeling is a very convenient computational approach to study the effect of these local changes on the potential 3-D binding patterns [3].

In addition to the effect of differences in the amino acid residues in the target, variations in the length of a monomer used as a ligand in dentin bonding, and the conformational flexibility of these molecules may also play a role on the binding process. This is because ligand binding is favored by the ability of the ligand to adapt favorable conformations to bind to available target sites. A well-established hydrophilic monomer used in dentin bonding is a methacrylate phosphate of alkane diol known as MDP (10-methacryloyloxydecamethylene phosphoric acid). This primer, with ten methylene spacer groups in its structure, has been recognized as an efficient self-etching primer. A second similar monomer MHP (6-methacryloyloxy-hexamethylene phosphoric acid) with six spacer groups is currently under investigation as another promising primer monomer for dentin bonding. Some changes in collagen-ligand aggregate binding behavior have been reported as a function of spacer groups in some primer monomers. Nishiyama *et al* [10] reported, for example, that the incorporation of incremental number of -CH<sub>2</sub>- spacer groups in a N-methacryloyl-omega-amino acid primer caused changes in adhesive behavior to dentin. The underlying cause or causes of such possible changes in the number of spacer groups on bond strength has not been previously examined. There are no previous computer-modeling reports to analyze the potential role of the number of spacer groups in modifying collagen-ligand interactions. Since the total number of spacer groups defines the extent of conformational freedom by increasing torsional centers, changes in the number of spacer groups facilitates improved ligand flexibility. Consequently, changes in the aggregate binding properties may be modified by the ability of the ligand molecules to adapt changes in torsion angles to align the ligand recognition sites to the complementary binding sites of the target. The recognition sites of the ligand molecules are typically functional groups such as carbonyl groups, phosphoric acid, carboxylic acid etc. The number of spacer groups in the ligand may alter the ability of the bound ligand molecule to stabilize in bound conformational states where either (or both) of non-bonded and electrostatic contributions to ligand-target interaction energy may be enhanced by optimized alignment and adaptation of the ligand molecule to the target [11]. Therefore, an understanding of the role of incremental spacer groups in determining collagen-ligand interaction energy may help design optimum dentin priming formulations. The objective of this investigation was to characterize the role of limited differences in amino acid residue sequences in the collagen target structure as well as the role of conformational freedom of a relatively small dentin primer

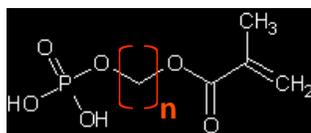
monomer modified with different levels of spacer groups on ligand binding to the collagen target.

## MATERIALS AND METHODS

### Direct (Target-Based) Docking Methods

**Collagen Target Structure:** Three homologues of type 1 collagen molecule with potentially important local differences in molecular surface topography and charge centers were used as collagen structural models. These homologues are designated as CGL, CGD and QSU. The typical (GLY-PRO-HYP)<sub>10</sub> sequence was designated as the CGL. The rationale for the selection of this model in the study is that it represents the most prevalent residue sequence in type 1 collagen and provides a stable triple helix. The presence of -GLY- residue in the triple helix is of importance in the stability of the triple helix motif in collagen structure. CGD is designed with the amino acid residue sequence of (PRO-HYP-GLY)<sub>4</sub>-(PRO-HYP-ALA)-(PRO-HYP-GLY)<sub>5</sub> and was selected because it has been synthesized and crystallographically characterized by Bella *et al.* [12] as an important triple helix motif with a surrounding hydration structure. This structure was extracted from RCSB PDB [13] file 1CGD. Bella *et al.* showed that ALA substitution for GLY in the (PRO-HYP-GLY) sequence causes a slight expansion of the triple helix locally at the site of substitution. The residue sequence in QSU is given by (PRO-HYP-GLY)<sub>4</sub>-(GLU-LYS-GLY)-(PRO-HYP-GLY)<sub>5</sub>. The QSU model was synthesized and crystallographically characterized by Kramer *et al.* [14] and is designated as 1QSU in the RCSB PDB. From the above sequence, it is evident that the QSU model contains a single pair of GLU and LYS residues (substituted for PRO and HYP residues) in the middle of the (PRO-HYP-GLY)<sub>10</sub> sequence. Kramer *et al.* [14] have shown that these residues alter local charge distributions. The localized spatial and charge differences resulting from such local changes in the repeat structure may be important in determining the ligand-collagen interaction energy. The selection of all three of the above models in the target-based simulations was intended to investigate the effect of these local changes in the target structure on the interaction energy.

**Ligand Structures:** The ligand structures studied were also based on a series of methacrylate phosphates of an alkane diol, derived by modifying the number of methylene spacer groups in the popular dentin primer monomer MDP (10-methacryloyloxydecamethylene phosphoric acid). The series can be represented chemically as:



The series were designed with differences in the number of methylene spacer groups to confer different levels of conformational degrees of freedom. The number of methylene spacer groups (*n*) varied from 2 to 12 with increments of two spacer groups between successive molecular structures in the direct docking studies. The resulting chemical compounds can also be described (using the nomenclature used for MDP in dentin adhesives) as 2-methacryloyloxy-dimethylene-

phosphoric acid, 4-methacryloyloxytetramethylene phosphoric acid, 6-methacryloyloxyhexamethylene phosphoric acid, 8-methacryloyloxyoctamethylene phosphoric acid, 10-methacryloyloxydecamethylene phosphoric acid and 12-methacryloyloxydodecamethylene phosphoric acid. Such ligand structures are identified in this study as [TWO], [FOU], [SIX], [EIG], [TEN], [TWE] etc., to highlight their differences in the number of spacer groups. Increase in the number of methylene groups in ligand structure introduces increasing torsional (conformational) degrees of freedom at each added single bond due to the spacer groups. The number of rotatable bonds increased from 8 to 18 as a result of the incorporation of increase in spacer groups from two to twelve in the structure. Docking studies were modeled with ligand molecules in their protonated states. 3-D structures of each of the above ligand molecules were built using Sybyl molecular modeling software (version 7.1.<sup>9</sup>) The structures were then subjected to energy minimization (conjugate gradient method) to obtain a nearest minimum energy conformational state for each ligand structure. The interactions between these ligands and the different collagen target structures were assessed by docking simulations of the ligand conformations to a rigid and static type 1 collagen molecule confined within a grid box using AutoDock software version 3.05<sup>f</sup>. Lamarckian Genetic Algorithm method in AutoDock was used for docking simulations [15]. The grid used in AutoDock was 120 x 120 x 120 along x, y, and z axes with a spacing of 0.375 Å. This facilitated a grid-based approach of interaction energy calculations accommodating the large collagen target molecule. During the docking simulations, the ligand molecules are subjected to translational and orientational changes by x, y, z axial translations and angular rotations of the molecules in 3-D space within the grid box. The dihedral angles at torsional centers are also varied. Interaction energy calculations are simplified using AutoGrid<sup>f</sup> with pre-calculated interaction energy values for non-bonded and electrostatic interactions for specific atom or charge probes of interest, if located at each grid point. The interaction energy at any point of the probe of interest is determined by trilinear interpolation and summed over all atoms and charges involved. Hydrogen bonding is included in the calculated non-bonded energy. Interaction energy in AutoDock (version 3.05) is calculated as docking and binding energy. Docking energy is the algebraic sum of non-bonded and electrostatic energy with the addition of internal free energy of the ligand. Binding energy represents the algebraic sum of the non-bonded and electrostatic energies with the addition of a positive torsional free energy term associated with ligand torsional freedom. The lowest docking and binding energy values of 50 bound conformational states were calculated and analyzed. The changes in docking/binding energy values due to the increased number of spacer groups in the different molecules were also analyzed. Changes in interaction energy range and mean for each ligand vis-à-vis each collagen target were assumed to result from changes in the bound conformational states of a ligand as a result of changes in the number of spacer groups.

<sup>9</sup> Tripos, St. Louis, MO

<sup>f</sup> Automatic Docking of flexible ligands to Receptor, AutoDock User Guide version 3.05, Scripps Research institute, La Jolla, CA.

The calculated energy values were analyzed by analysis of variance (ANOVA) and Tukey multiple comparisons using JMP statistical software.♦ Two-way ANOVA procedure was used to analyze the changes in computed docking and binding energy values due to modifications to collagen structural model and changes in the number of spacer groups in the ligand. In addition, the effect of these changes to collagen and ligand on the electrostatic and van der Waals energy contributions to docking/binding energy values was also analyzed. The analysis was focused on the effect of independent variables (collagen model and ligand differences) on the dependent variables (docking energy, binding energy) as well as the separate van der Waals and electrostatic contributions to these energies, computed by AutoDock program. Statistically significant differences in the computed energy mean values were assessed at  $\geq 95\%$  confidence level ( $P \leq 0.05$ ) using Tukey multiple comparisons.

### Indirect (Ligand-Based) Method

The indirect method used in this study was a pharmacophore search using the HipHop algorithm of Catalyst (4.11) software.⁵ HipHop generates feature or function-based pharmacophore hypotheses [15]. In the function-based representation, atoms are grouped into functional features that are important in ligand-target binding. Typical functionalities used are hydrogen bond acceptor, hydrogen bond donor, acid functionality (negative ionizable function at physiological pH 7), base (positive ionizable function at physiological pH 7), aromatic ring and hydrophobic group. Ligands are treated in their respective protonated states. In the HipHop algorithm, a pharmacophore model or hypothesis is generated as a distinct three-dimensional configuration of chemical functions located at centers of tolerance spheres [4, 5, 16]. A tolerance sphere defines that area in space that should be occupied by a specific type of chemical functionality in the bound ligand structure. Hydrogen bond donors and acceptors also include a vector to incorporate the constraint of bond directionality. Misalignment of any functionality from its reference sphere of tolerance or directional constraint can be expected to weaken the interactions between the chosen functionality and its complementary functional site in the target. Pharmacophore modeling and analysis were carried out on a SGI workstation (Octane, MIPS with R12000) running IRIX 6.5.28m operating system with a 400 MHz processor with 1536 MB of memory. All structures of primer molecules included in the study were generated in Sybyl and imported into Discovery Studio 1.7⁵ to create a trial set. Molecular flexibility was included by treating each molecule as an ensemble of conformers representing different accessible areas in the 3-D space. Confirm module was used to generate conformers using Discovery Studio 1.7 running on a Linux server [Fedora Core 6 operating system with Intel (R) Xeon CPU 3 GHz]. Confirm uses poling algorithm to select "best" conformers by promoting diversity in conformational selection [17]. The maximum number of conformations opted for consideration was 500, and the energy window of the conformational states considered was set to 20 kcal/mol above

the minimum energy conformational state in the ensemble. Based on the adhesive characteristics of the monomers with six and ten methylene groups (respectively known as MHP and MDP, as indicated earlier), these monomers were treated as reference molecules. In HipHop modeling, potential pharmacophore models of common functional alignment of a set of bioactive ligands are determined by a stepwise procedure [4, 5, 16]. The first step involves data input of the ligand structures. We studied eleven structures with unit incremental difference in number of methylene spacer groups in methacrylate phosphate series of alkane diol (see Ligand Structures above). A conformational search is used to identify low energy conformations within a user specified energy window. This is followed by the extraction of the functions of the ligands to build function based structural models of ligands. The program finally generates user specified number (set to 10 in this study) of pharmacophore hypotheses using the common alignment patterns of the functional placements in all ligands included in the search, and scores the different hypotheses from the highest to lowest probable ranks. The highest ranked hypothesis (01, Table 1) was used for detailed analysis of individual ligands. The ligands were mapped to the pharmacophore and assessed visually, and by parameters such as the fit value, conformational energy, etc., computed by the program as well as by the number of conformations of each ligand used in generating the pharmacophore model. Bound conformations (top-ranked docking poses) of selected ligands from AutoDock were also mapped to pharmacophore hypothesis 01 in a rigid manner to ascertain whether bound conformational poses generated using a target in the simulations correspond to pharmacophore patterns generated using ligand-based modeling with no target.

## RESULTS

### Direct Docking Studies

Figs. (1) shows the CGL (a), CGD (b) and QSU (c) variants of the collagen structural models used in the study. The initial geometry-optimized, energy-minimized structures of methacrylate phosphate series of alkane diol containing 2, 4, 6, 8, 10 and 12 methylene spacer groups are shown in Fig. (2). The number of active torsional centers in the respective structures were 8, 10, 12, 14, 16 and 18, respectively. The starting structures had significant differences in their geometric configurations, especially in their tendency to fold with increasing spacer groups. Fig. (3) illustrates docked conformations of ligand [TEN]. Analysis of docked ligand conformations revealed a distinctly evident change from folded initial conformations to extended bound conformational states. Such changes from their starting configurations through torsional changes at single bonds promote stable binding interactions such as hydrogen bonding, steric and charge interactions etc. Fig. (4) illustrates potential hydrogen bonds between the target and the docked ligand. Other key features of collagen-ligand binding (such as electrostatic and steric complementarity) have been reported in our previous publications [3, 7].

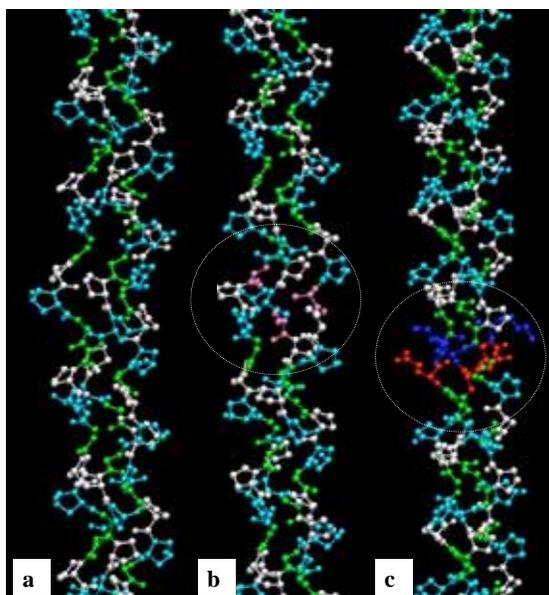
Cluster analysis of docked conformations at rmsd value of 2 Å revealed multiple clusters for each ligand, and differences were observed between cluster distributions when different ligands were evaluated. A visualization of selected

♦ JMP® Statistical Discovery Software, SAS Institute, Cary, NC

⁵ Accelrys, 9685 Scranton Road, San Diego, CA 92121

⁵ Accelrys, 9685 Scranton Road, San Diego, CA 92121

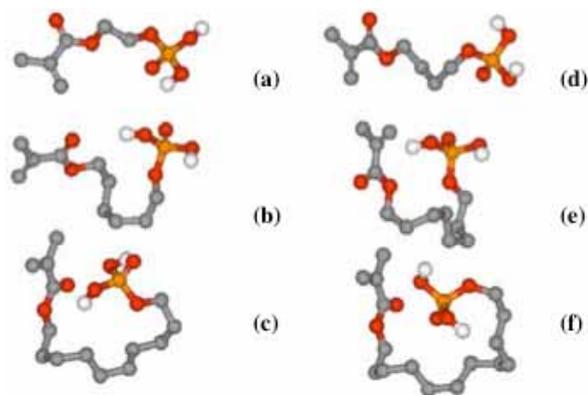
conformations of ligand [TEN] in Fig. 3 indicates that the different cluster groups appear to be present at different binding pockets on the target. Multiple binding pockets occur because of repeating triplets (such as -GLY-PRO-HYP- in CGL model) along the length of the collagen molecule. It is to be pointed out that because of the repeating amino acid triplet sequences in a triple helix, most binding sites are very similar along the length of the collagen molecule. Multiple clusters indicate that different target sites where docking occurred are energetically favored by the conformations in the respective clusters. Because of the similarity of the target sites, multiple cluster formations are clearly favored. In our isothermal titration calorimetric (ITC) experiments between acid soluble type 1 collagen and 2-methacryloyloxydimethylene phosphoric acid (compound [TWO] in this study), no binding saturation was observed, indicating multiple binding sites on the target.



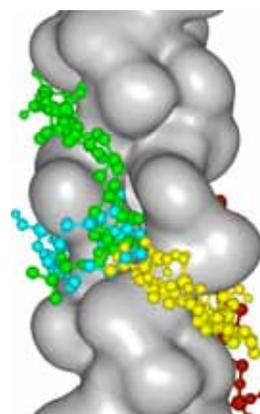
**Fig. (1).** Three representative structural variants of type 1 collagen studied; limited differences in amino acid residue sequences are included within the dotted circles in (b) and (c). (a) CGL with (GLY-PRO-HYP)<sub>10</sub> sequence (b) CGD with (GLY-PRO-HYP)<sub>10</sub> sequence having -ALA- substitution for -GLY- in the fifth triplet (c) QSU with (GLY-PRO-HYP)<sub>10</sub> sequence having -GLU-LYS- substituted for -PRO-HYP- in the fifth triplet.

AutoDock search yielded docked conformations with individual docking energy values (kcal/mole) ranging from -3.22 to -8.21, with least square (LS) means ranging from -4 for QSU to -6.2 for CGD. The individual binding energy values ranged from +0.94 to -3.45 kcal/mole with LS means ranging from -0.45 for CGL to -2.37 kcal/mole for CGD. Only a relatively small fraction of the cases (32 out of 900 conformations) showed positive binding energy values unfavorable for binding. Thus computational analysis indicates that most of the binding conformations potentially survive and contribute to ligand adhesion to the target.

Two-way ANOVA results of the interaction energy values (docked energy, binding energy) and the electrostatic and van der Waals contributions to the docking/binding energy

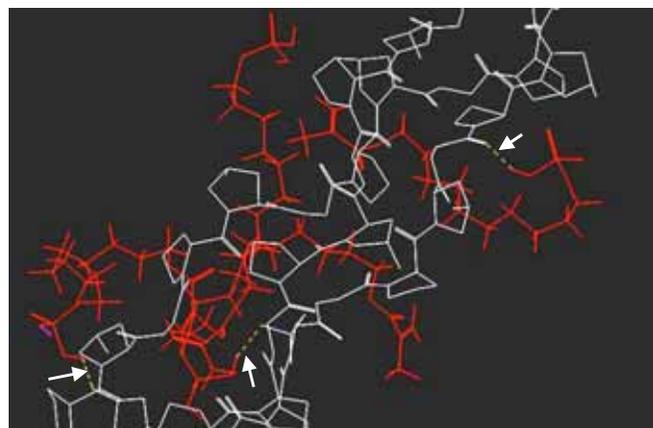


**Fig. (2).** Illustrations of six monomer ligands with even number of -CH<sub>2</sub>- spacer groups from 2 to 12: (a) Two (b) Four (c) Six (d) Eight (e) Ten (f) Twelve. Nonpolar hydrogen atoms are not shown.

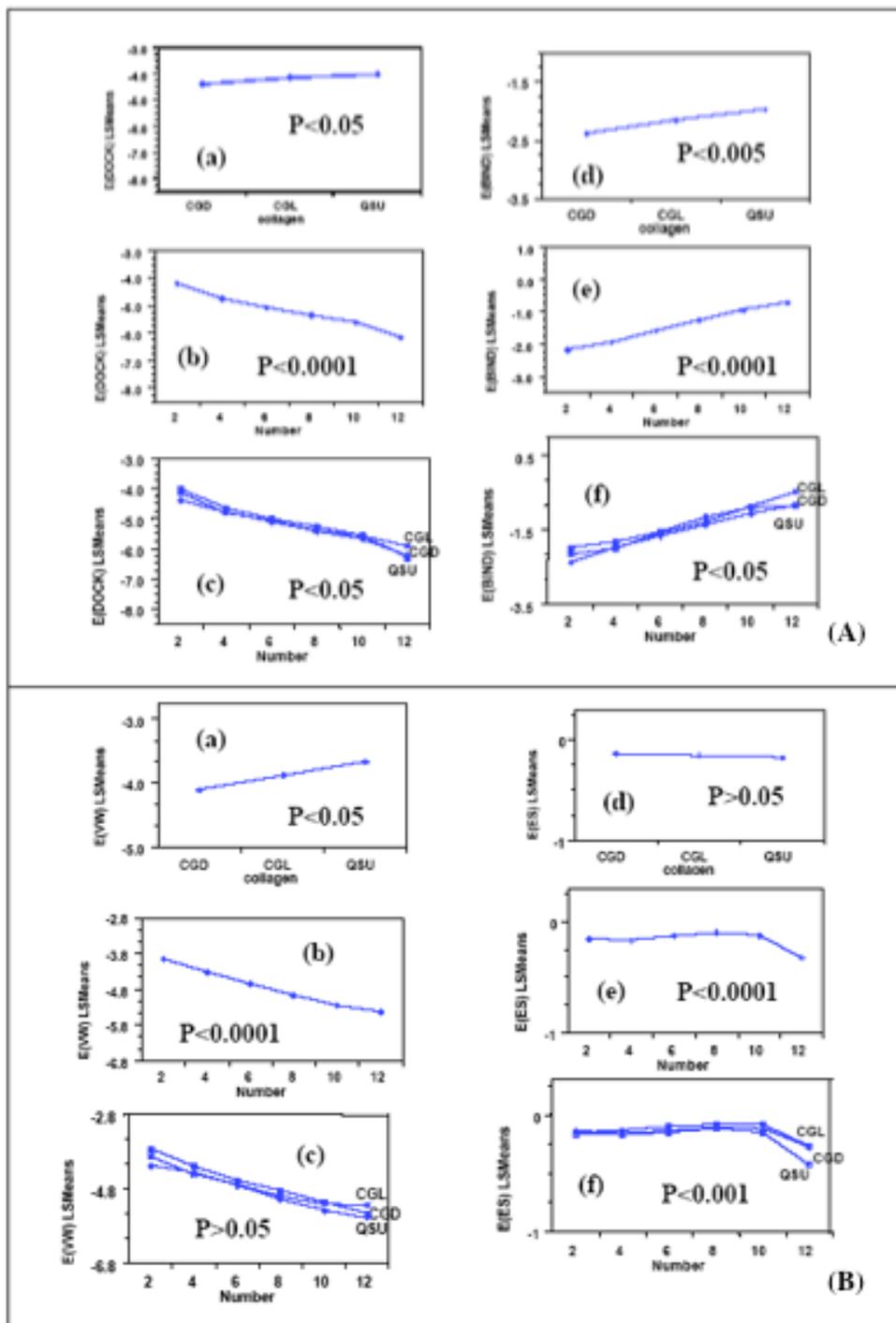


**Fig. (3).** Representative illustrations of docked conformations: Visualizations of docked ligand with  $n = 10$  CH<sub>2</sub> groups on the QSU collagen structural model. Note clusters of conformations at different sites along the cavity tracks.

were also examined as a function of the main independent variables (collagen and the number of -CH<sub>2</sub>- groups and their interactions). Profile plots of the docking and binding energy values reveal statistically significant main and interactive effects ( $P < 0.05$ ), as seen in Figs. (5A(a,b,c)) and (5A(d,e,f)).



**Fig. (4).** Illustration of docked ligands (red) on collagen target (white). Note binding at multiple sites and the predicted hydrogen bonds (see dashed lines highlighted by arrows).



**Fig. (5).** Profile plots of interaction energy (docking and binding energy) and the van der Waals and electrostatic contributions to docking/binding energy values. Main and interactive effects of independent variables (Collagen, number of spacer groups) on docking energy {A(a,b,c)}, binding energy {A(d,e,f)}, and the corresponding van der Waals contribution {B(a,b,c)} and electrostatic contribution {B(d,e,f)}.

The effect of changes in the number of spacer groups on the docking/binding energy values was highly significant [Fig. 5A(b),  $p < 0.0001$ ], but changes in collagen composition were found to result in only marginal changes ([Fig. 5A(a),  $p < 0.05$ ] in interaction energy. In addition, the docking energy decreased with increasing number of methylene groups [Fig. 5A(b)], but the binding energy tended to increase [Fig. 5A(e)] because of positive torsional free energy contribution

due to increased torsional freedom. Similarly analysis of the electrostatic and van der Waals energy contributions to the ligand-collagen interaction energy in Figs. 5B(a,b,c,d,e and f) also showed statistically significant differences as a function of the number of spacer groups ( $p < 0.0001$ ), but changes in electrostatic energy with collagen composition [Fig. 5B(d)] was not significant ( $p > 0.05$ ).

**Table 1. List and Features of Pharmacophore Hypotheses**

Hypotheses	Features*	Rank
01	NHHA	96.3234
02	NHHA	94.0272
03	NHHA	90.3847
04	NHA	89.8460
05	NHHA	89.7677
06	NHA	89.4091
07	NHHA	88.5373
08	NHHA	87.0065
09	NHHA	86.6241
10	NHHA	86.5834

\*N= Negative ionizable function, H=Hydrophobic function and A= hydrogen bond acceptor

**Table 2. Details of Compound Name (Column 1), Input Parameters\* [Columns 2-4] Used in HipHop Calculation, and the Computed Output Data\*\*(Columns 5-6)**

Name	Principal	MaxOmitFeat (MOF)	Confs	Conformational Energy	Best fit
TWO	1	1	195	16.832	2.870
THR	1	1	203	0.525	2.093
FOU	1	1	300	7.574	3.140
FIV	1	1	300	10.291	3.703
SIX	2	0	432	14.178	3.880
SEV	1	1	374	5.993	3.970
EIG	1	1	345	12.511	3.973
NIN	1	1	330	8.156	3.987
TEN	2	0	337	4.372	3.999
ELE	1	1	405	18.100	3.986
TWE	1	1	410	17.852	3.996

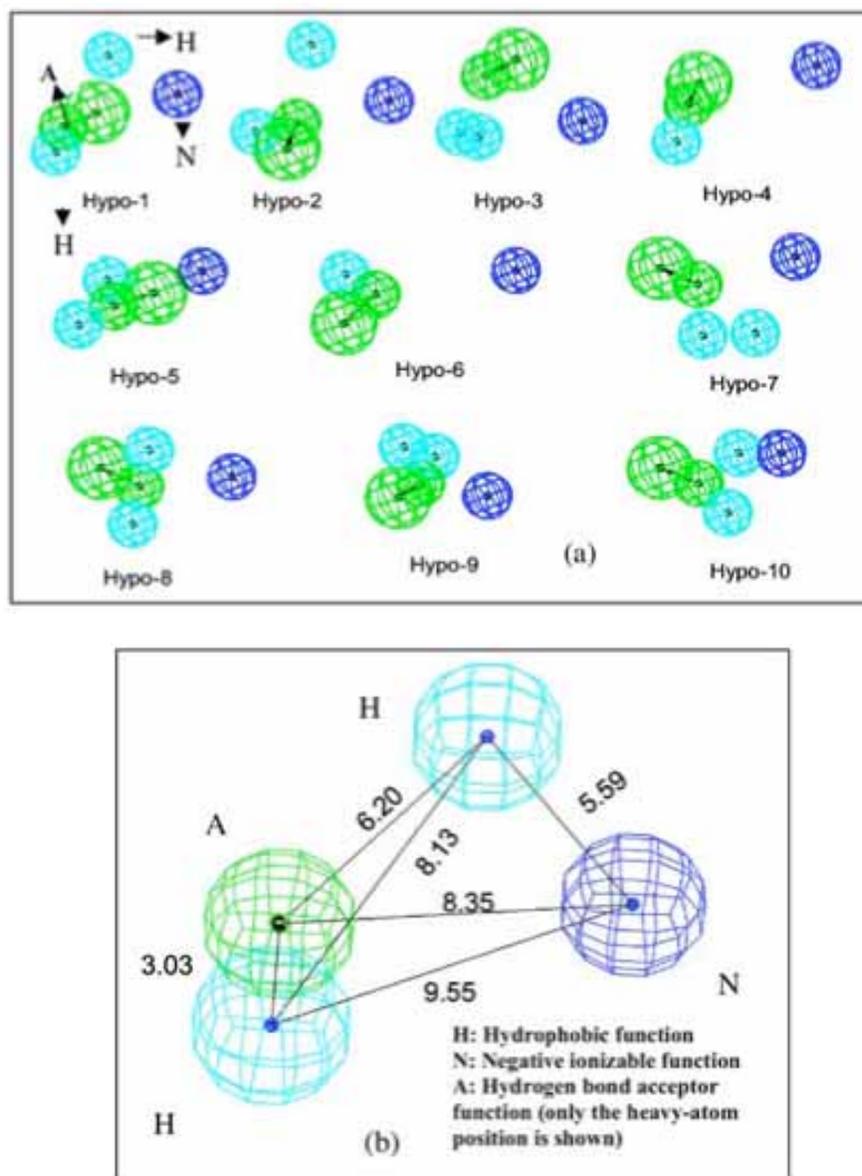
\*includes the input values of columns 2 and 3, and the number of conformations (out of 500 maximum allowed) used in generating the pharmacophore, listed in column 4.

\*\*Conformational energy and fit values {best fit, MOF=0 except for [TWO], MOF=1 for [TWO]} were computed using Compare/Fit option in Catalyst for hypothesis 01 in Table 1)

### Indirect Pharmacophore Search

Fig. (6(a)) shows ten pharmacophore hypotheses generated in the indirect approach. Table 1 lists the different hypothesis combinations with details such as the functional features represented in the pharmacophore and the order of ranking (based on the criteria that a higher ranking number indicates a lesser likelihood for the hypothesis to be generated by chance). Clustering the top 10 hypotheses using Hierarchical Average Linkage method at the two-cluster level showed that hypotheses 04 and 06 belong to a different cluster than the rest of the models (see Table 2). All the hypotheses except 04 and 06 have a common functional combination

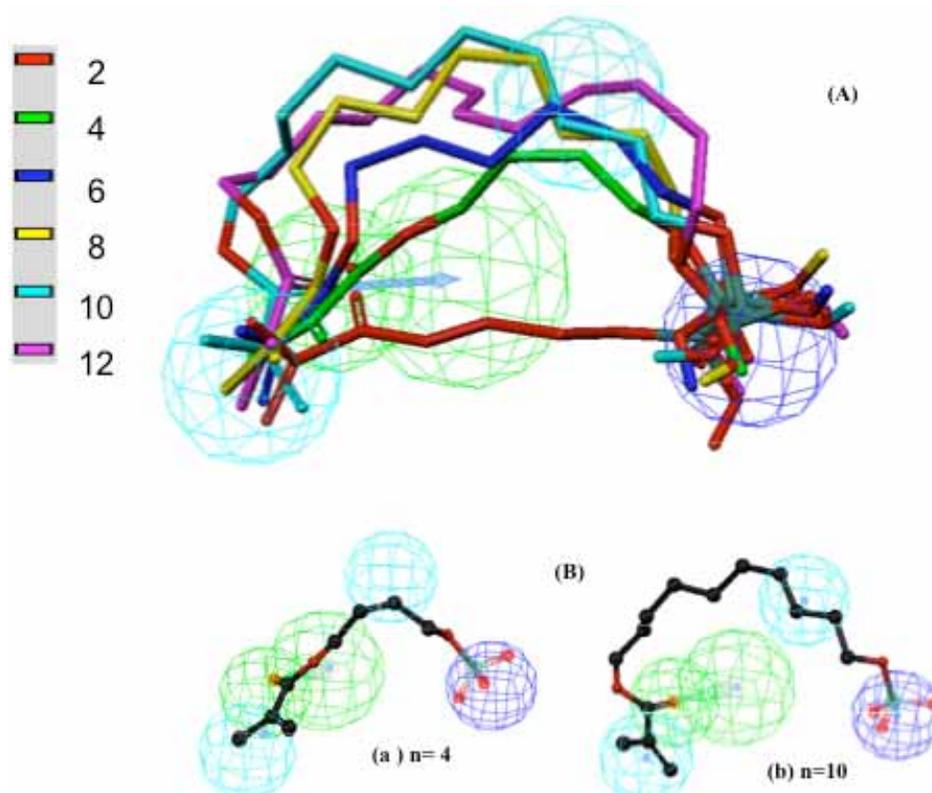
of NHHA in the pharmacophore model, where N represents a negative ionizable function, HH indicate two hydrophobic groups, and A, the hydrogen bond acceptor group. The functional combination for 04 and 06 hypotheses is NHA. The ranking number for the hypotheses ranged between 96.32 and 86.58, which is relatively narrow, but for the purpose of illustration we selected 01 hypothesis (which also had the highest ranking) for all our subsequent mapping and analysis. Table 2 shows the list of ligands with the respective details including (a) maximum number of features omitted [i.e., MaxOmitFeat or (MOF) = 0 for none and 1 for one], (b) the number of monomer conformations used in pharmacophore



**Fig. (6).** (a) Visualization of top ten Pharmacophore models by indirect modeling simulations. All hypotheses except 04 and 06 contain NHHA functional grouping where N is a negative ionizable function, HH are two hydrophobic groups and A is a hydrogen bond acceptor function, as illustrated for hypothesis 01. 04 and 06 hypotheses were made up of NHA functional grouping. (b) The distance matrix of the 3-D functional placements in the highest ranked pharmacophore model. Only the heavy atom is shown, not the bond direction vector.

modeling and (c) the computed best “fit” values of the best conformation of each of the respective ligands vis-à-vis the pharmacophore hypothesis 01. Monomer with six spacer groups used the highest number of conformations, but the highest fit was for the monomer with ten spacer groups. Both of these monomers were used as the reference molecules in the pharmacophore model search, and received higher weighting in the search. Based on consideration of the number of conformations and fit values as criteria for ligand ability to bind to the target, it is reasonable to conclude that the ligands with five or more spacer groups in the series investigated are likely to have improved binding to the intended target. Ligand overlay with the pharmacophore revealed that phosphoric acid is the negative ionizable function, and that the hydrogen bond acceptor corresponds to the carbonyl

group of the methacrylate functionality. Two hydrophobic groups (HH) include (a) the acrylic group in the methacrylate function and (b) the spacer groups. Fig. (6(b)) shows a distance matrix of the functional placements in the highest ranked hypothesis 01. Note that the functional placements are defined by an optimum center and range of tolerance values in space. An overlay of monomers with different numbers of CH<sub>2</sub> groups on hypothesis 01 model is shown in Fig. (7A). An interesting feature of the overlay is that all the ligands except the monomer with only two CH<sub>2</sub> spacer groups have all the respective functions positioned within the tolerance spheres of corresponding functions of the pharmacophore model. The monomer with only two CH<sub>2</sub> groups does not align to hydrophobic group corresponding to the spacer links.



**Fig. (7).** (A). Mapping (best fit) of individual ligands with even number of spacer groups into the highest ranked Pharmacophore 01 hypothesis 3-D model. Note the ability of monomers with different number of -CH<sub>2</sub>- groups to align their functional placements to match the 3-D placement of pharmacophore tolerance spheres of the respective functions through conformational flexibility associated with the spacer groups. See table 2 for computational details. (B). Visualization of ligand mapping (best fit) to the pharmacophore to illustrate how torsional freedom at spacer groups renders conformational flexibility to align the functions to a common pharmacophore geometry of hypothesis 01. (a) n=4; (b) n=10. Note that hydrogen bonds are not shown for better clarity. See table 2 for computational details.

Figs. (7B(a,b)) also illustrate the contrasting alignment of selected ligands to the pharmacophore. The torsional freedom of the rotatable single bonds provide the necessary molecular flexibility to facilitate multi-spacer monomers to assume folded or extended conformational states to adjust their feature positions to match the distance matrix of functional placements defined by the pharmacophore. Thus conformational flexibility appears to provide a mechanism to align ligand functionalities to the complementary functional sites of the (hypothetical) target. Important parameters such as fit, the number of conformations and the lowest energy of the hypothetical bound conformational state of the ligand may help define the ability of the ligand to bind to the target. These values in the indirect method together with the interaction energy values in the direct docking studies may be valuable parameters for virtual screening of individual primer monomers.

## DISCUSSION OF RESULTS

In the direct method used in our initial study, the effect of collagen compositional and ligand conformational changes on the computed ligand-collagen interaction energy was estimated using an automatic docking algorithm. The results revealed significant main and interactive effects of these changes on the intermolecular interactions. These results were analyzed using the current generation of the pharma-

cophore model, (c) the conformational energy of the best fit conformer concepts of adhesion of ligands to a biological target molecule, where the 3-D placement of functionalities on the ligand plays a central role in ligand binding to a target binding site. In particular, the steric and electrostatic complementarities of the ligand recognition sites and the target binding site or sites may cause differences in non-bonded and electrostatic contributions to the intermolecular interaction energy, and these differences are estimated in docking simulations.

The primary compositional changes evaluated in collagen were substitutions of amino acid residues in a typical (PRO-HYP-GLY) residue sequence of type 1 collagen  $\alpha$ -chains. Two modifications assessed in this study were substitutions of ALA for GLY in one of the variants, and GLU and LYS for PRO and HYP, respectively, in the other. These changes had a marginally statistically significant effect on the computed mean docking/binding energy values [ $p < 0.05$ , Figs. (5A(a)) and (5A(d))]. It is well known that the presence of GLY at every third amino acid position in the -GLY-X-Y-repeating sequence in each of the three  $\alpha$ -chains of collagen molecule is essential because a larger amino acid would not fit into the restricted space at the center of the triple helix where the three chains come together. Substitution of a single -ALA- for -GLY- at the center of a collagen-like (PRO-HYP-GLY)<sub>10</sub> peptide has been shown to result in a small

local untwisting of the triple helix at the substituted site, causing a small expansion of the molecule. Despite the local distortion, the triple helix motif was demonstrated to be stable, although a decrease in thermal stability was observed (Bella *et al*<sup>12</sup>). Visualization of the computed docking results showed that docking of ligands would be favored to occur at these locally expanded sites due to favorable steric conditions. Typically this was found to show a small but statistically significant increase ( $p < 0.05$ ) in the van der Waals energy contribution to docking energy in the CGD structure [Fig. (5B(a))]. GLU-LYS substitution for PRO-HYP in the middle of (PRO-HYP-GLY)<sub>10</sub> sequence in the Kramer model showed changes in local electrostatic charge distribution profile in the collagen molecule. This was found to have no significant effect on the electrostatic contribution to the docked energy [ $p > 0.05$ , Fig (5B(d))]. This appears to be the result of absence of charge contrasts in the ligand and collagen structures. Increase in the number of spacer groups in the ligand also typically showed a highly significant change in the docking energy to more negative values favoring improved collagen-ligand interaction [Fig. (5A(b))]. Increasing the number of spacer groups increases the conformational flexibility of the molecule due to an increase in the number of rotatable bonds in the ligand structure. Because of this increased torsional freedom, the molecule has the flexibility to find more optimum feature coordinate positions with respect to the target feature positions to enhance for -GLY- at the center of a collagen-like (PRO-HYP-GLY)<sub>10</sub> peptide has been shown to result in a small local untwisting of the triple helix at the substituted site, causing a local spatial for -This appears to be the result of absence of charge contrasts in the ligand and collagen structures.

Non-bonded and/or electrostatic interaction energy contributions to docked energy. However, while the docking energy decreased to more negative values, the binding energy tended to increase to less negative values [See Figs. (5A(b)) and (5A(e))]. This is expected because the binding energy term in AutoDock includes a positive torsional free energy term for each added torsional degree of freedom. This would indicate that the number of spacer groups needs optimization for most efficient priming.

It is to be noted that the molecules with multiple rotatable bonds often assume extended conformations in the docked state vis-à-vis a folded geometry in the starting state (see Figs. (2) and (3)). Ligand-target Interaction in an extended conformational state of the ligand may also enhance interaction energy through binding interactions with more target atoms or sites,

Differences in ligand-collagen combination subsets [Figs. (5A(c)) and (5B(f))] showed a significant interaction effect on both docking energy ( $p < 0.05$ ) and electrostatic energy ( $p < 0.001$ ). Such interaction effects are expected because differences in the length and flexibility of the ligand can help bring the complementary sites of the ligand and target closer or farther apart depending upon the alignment of the ligand and collagen functional features in the docked complexes.

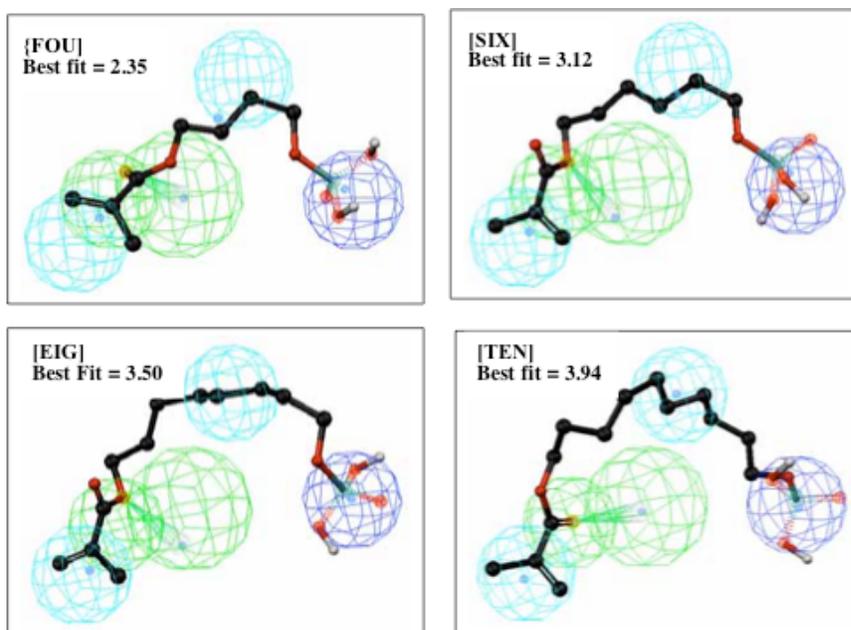
The pharmacophore hypotheses generated in our indirect modeling give additional support to the role of conformational flexibility of the molecule in ligand binding. The pharmacophoric model generates a potential geometry of the

binding conformations of the ligands to the target [4, 5], i.e., the model presents a likely snapshot of the spatial positions of functional features of the ligand in alignment to the complementary functional features of the target. Conformational flexibility of a ligand may enable it to assume a geometric configuration to arrive at a more favorable functional alignment vis-à-vis target functionalities by the rotation of the single bonds at the location of the spacer groups. Fig. (7B(b)) shows the alignment of selected molecules (with 4 and 10 spacer groups) to the pharmacophore hypothesis 01, and illustrates how the flexibility of the single bonds in the ligand facilitated their match to the pharmacophore model.

Both direct and indirect simulations use low energy conformations to assess binding interactions. In both methods, the ligands were treated in their protonated states to determine different ligand binding poses. AutoDock calculations use a target structure, but HipHop determines binding poses using positions of functions of ligands assumed to be in the bound conformational states with respect to an assumed target. Thus the two methods differ only with respect to the use of the target in the simulations, but they nevertheless identify the ligand binding poses from low energy conformations under similar criteria. For this reason, both simulations should arrive at similar and mutually consistent ligand binding poses. Our results showed that the binding poses generated by target-based simulations were satisfactorily mapped to one or more pharmacophore models generated by ligand-based HipHop simulations. For example, Fig. (8) shows illustrations of selected top-ranked poses from direct docking studies of ligands with six, eight and ten spacer groups mapped to the pharmacophore hypothesis 01. The poses were mapped in a rigid manner. The fit of the best-matched poses are shown. The fit values appear to improve with improved conformational flexibility. The results elegantly demonstrate that both direct and indirect simulations lead us to reasonably matching binding poses.

As pointed out earlier, the collagen molecule contains multiple binding pockets, as is evident from Figs. (3 and 4). However, these pockets are otherwise essentially similar because of the repeating amino acid triplet sequences. For example, the three hydrogen bonds in Fig. (4) were all between hydroxyproline of collagen and phosphoric acid of MDP, despite the fact that the ligands were docked at different target locations. Consequently the binding poses at similar sites at different locations along the cavity tracks are also often similar. This explains why one or more pharmacophore hypotheses reasonably match to docking poses of ligands at different target sites in target-based simulations. Differences in binding site location appear to have only limited effect on binding poses. This is a very favorable circumstance for improved overall adhesion, because more binding pockets ensure more efficient priming of the entire target.

The computational approach used in this study needs additional experimental support for validation. Nishiyama *et al*<sup>10</sup> have shown a significant effect on the measured bond strength between dentin and restoration when N-methacryloyl-omega-amino acid primer monomers with differences in the number of -CH<sub>2</sub>- spacer groups were used to prime dentin. Our results give a valid theoretical basis for these observed results. In addition, although the phosphoric



**Fig. (8).** Illustration of bound conformations (top-ranked poses) of ligands [FOU], [SIX], [EIG] and [TEN] generated in AutoDock mapped to pharmacophore hypothesis 01.

acid function in MDP has been shown to interact strongly with hydroxyapatite (HAP), MDP-bonded restorations show higher bond strength to collagen-HAP composite structure of dentin than to HAP-rich enamel. Our results suggest that this may result from efficient priming of the dentinal collagen by MDP, which also helps bonding to the HAP phase of dentin. Thus, overall adhesion to dentin can potentially be strengthened by promoting simultaneously improved non-bonded and electrostatic primer interactions to dentinal collagen as shown in this study, while improving chemical adhesion of the primer to the HAP phase of dentin as well. Together, these bonds ensure that both HAP phase and dentinal collagen are properly bonded in the hybrid layer to reduce or eliminate relatively weak interfaces between the participating phases as far as possible. It is well known that the weak links or interfaces in a two- or multi-phase structure are the pathways of mechanical and environmental failure.

#### ACKNOWLEDGMENTS

This study was supported by a grant (R01 DE 14370) from the National Institute of Dental Research of the National Institutes of Health, Bethesda, MD.

#### REFERENCES

- [1] Fischer E.: Ueber die Optischen Isomeren des Traubenzuckers, der Gluconsaure und der Zuckersaure. *Ber. Dtsch. Chem. Ges.* 23(2), 2611, (1890).
- [2] Koshland D.E.: Application of a theory of enzyme specificity to protein synthesis. *Proc Natl. Acad. Sci. USA.* 44(2), 98, (1958).
- [3] Vaidyanathan J., Vaidyanathan T.K., Ramasubbu N., Ravichandran S.: A Computational and experimental analysis of ligand binding to type 1 collagen. *Current Computer aided Drug Design* 1(4), 397, (2005).
- [4] Dror O., Shulman-Peleg A., Nussinov R., Wolfson H.J.: Predicting Molecular Interactions in silico: I. A Guide to Pharmacophore Identification and its Applications to Drug Design. *Curr. Med. Chem.* 11(1), 71, (2004).
- [5] Guner O.F.: History and Evolution of the Pharmacophore Concept in Computer-Aided Drug Design. *Current Topics in Medicinal Chemistry* 2 (12), 1321, (2002).
- [6] Nakabayashi N., Pashley D.H.: Book title, In Hybridization of dental hard tissues. Quintessenz Publishing Co.: Chicago., (1998).
- [7] Vaidyanathan J., Kasinathan C., Vaidyanathan T.K.: Biomimetic Recognition and Immunochemical Binding Assay of collagen-ligand binding *J. Adhesive Dent.* 5(1), 7, (2003).
- [8] Vaidyanathan J., Vaidyanathan T.K., Yadav P., Linaras C.E.: Collagen-ligand Interaction: Computer visualization and analysis, *Biomaterials* 22(21), 2911, (2001).
- [9] Veis A, Schlueter R.H.: The macromolecule organization of dentin matrix collagen. 1. Characterization of dentin collagen. *Biochemistry.* 3(11), 1650, (1964).
- [10] Nishiyama N., Suzuki K., Takahashi K., Nemoto K.: The  $pK_a$  effects of the carboxylic acid in N-methacryloyl-omega-amino acid on the demineralization and bond strengths to the teeth. *Biomaterials.* 25(23), 5441, (2004).
- [11] Vaidyanathan J., Vaidyanathan T.K., Kerrigan J.E.: Evaluation of intermolecular interactions of self-etch dentin adhesive primer molecules with type 1 collagen: Computer modeling and in vitro binding analysis. *Actabiomaterialia* 3(5), 705, (2007).

- [12] Bella J., Brodsky B., Berman, H.M.: Hydration structure of a collagen peptide. *Structure* 3(9), 893, (1995).
- [13] Berman H.M., Westbrook J., Feng Z., Gilliland G., Bhat T.N., Weissig H., Shindyalov I.N., Bourne PE.: The protein Data Bank. *Nucleic Acids Research*. 28(1), 235, (2000).
- [14] Kramer R.Z., Venugopal M.G., Bella J., Mayville P., Brodsky B., Berman H.M.: Staggered molecular packing in crystals of a collagen-like peptide with a single charged pair. *J. Mol. Biol.* 301(5), 1191, (2000)
- [15] Morris G.M., Goodsell D.S., Halliday R.S., Huey R., Hart W.E., Belew R.K., Olson A.J. Automated docking using a Lamarckian Genetic Algorithm and empirical binding energy function. *J. Comput. Chem.* 19(14), 1639, (1998)
- [16] Clement, O. Mehl A.T. "HipHop: Pharmacophores based on Multiple Common-Feature Alignments". In *Pharmacophore Perception, Development, and Use in Drug Design*. O.F. Guner, ed., International University Line, La Jolla, CA, 2000, pp 71-84.
- [17] Smellie A., Teig, S.L., Towbin P.. Poling: promoting conformational variation. *J. Comp. Chem.* 16 (2), 171, (1995).