Proposal of a 3D Peptide Pharmacophore of Muramyl Dipeptide-Type Immunostimulants. 2. Computer Docking to a Model Protein Binding Site†

Primož Pristovšek, Jurka Kidrič,* and Dušan Hadži

National Institute of Chemistry, Hajdrihova 19, P.O.B. 30, 1115 Ljubljana, Slovenia

Received February 10, 1997⊗

The conformation of the immunostimulant muramyl dipeptide (*N*-acetylmuramyl-L-Ala-D-iGln, MDP) selected by the application of the CCLUES method (preceding paper) as the best candidate for the bioactive conformation is closely related to one of its parent compound, *N*-acetylglucosaminyl-β1→4-*N*-acetylmuramyl-L-Ala-D-iGln-γ-diaminopimeloyl-D-Ala, when bound to the T4 lysozyme [Kuroki, R. *et al*. *Science* **1993**, *262*, 2030]. A series of active and inactive MDP analogues has been docked to the same binding site and analyzed for site-to-ligand group−group interactions. The docking experiments demonstrate that the binding site qualitatively discriminates between the diastereomers of MDP and between the active and inactive analogues. It therefore appears to be a suitable model of peptide binding to the putative receptor for immunostimulant MDP-type peptides. The conformation of MDP docked to the model binding site is taken for the *ab initio* calculation (3-21G basis set) of the molecular electrostatic potential that is representative of the 3D electrostatic pharmacophore of the peptide core. Also including the more general structure−activity relations it can be used as a starting point for the design of MDP mimetics.

INTRODUCTION

Considerations of the structure−activity relations of muramyl peptide analogues suggest that the immunostimulating activity is specifically connected with the L-Ala-D-Glu core of the molecules; only very limited variations in the amino acid type (e.g., Ala→Val, Glu→Gln), but none in the configurations (e.g., D-Glu→L-Glu leads to completely inactive analogues), are allowed,[1] while wide variations in the N- and C-terminals of the dipeptide fragment are acceptable. For the design of nonpeptidic mimetics it is, therefore, of importance to define the 3D pharmacophore based on the dipeptide core in the bioactive conformation of the lead peptide.

In the preceding paper[2] (hereafter referred to as the previous paper) we derived three most probable bioactive conformations of muramyl dipeptide (MDP, MurNAc-L-Ala-D-iGln, **I**, Figure 1; MurNAc: *N*-acetyl-muramic acid) by comparison of conformations of active and inactive analogues using the CCLUES[3] procedure; with the latter, the information encoded in the inactive diastereomers MurNAc-L-Ala-L-iGln (**II**), MurNAc-D-Ala-L-iGln (**III**), and MurNAc-D-Ala-D-iGln (**IV**, to a lesser extent) was most useful in the first step,[3] while in the second step for further selection analogues differing in molecular structure became necessary.[2] From the three final candidates the most probable one was selected by considering the accessibilities of target atoms. However, the suppositions, neglects, and simplifications of the CCLUES method[3,4] have to be considered; the results derived with the method should therefore be checked with an independent approach.

Amongst the conjectures inherent to the CCLUES method also is the selection of the groups that are believed to be of key importance for the interaction with the receptor site and hence for the distinction between active and inactive analogues. In the absence of structural information on the latter, investigation of binding mechanisms of MDP and related murein fragment derived peptides (MFDPs) to proteins is a possible source of support for the choice of critical contacts for the peptide fragment even if the protein function is different. Clearly, for such investigations reliable structural evidence of specific binding is the prerequisite. An attractive model binding site is that of the bacteriophage T4 lysozyme (T4L); this protein belongs to the group of v-type lysozymes[5] and contains a binding site for the L-Ala-D-Glu dipeptide connected to the muramyl ring at subsite D.[6]

Lysozymes of various types (c − hen egg white, g − goose egg white, v − bacteriophage) follow a general structure scheme; they consist of two domains embracing a cleft with the active site that accommodates six β(1→4) linked pyranoside units in subsites A−F. Some previous studies of binding based on crystallographic data[6−9] and computer simulations[10,11] of inhibitor−lysozyme complexes were focused on the interaction of the glycosidic units with the subsites with particular concern for the lytic mechanism. Our interest being centered upon the molecular determinants of the binding of MFDPs and particularly differentiation between the diastereomers of MDP, we found the v-type lysozymes to be most suited for this purpose. They differ from the c-type in that the substrate (consisting of alternating GlcNAc-β1→4-MurNAc units; GlcNAc: *N*-acetylglucosamine) must contain for full activity L-Ala-D-Glu (or iGln) attached to muramic acid;[12] this indicates the existence of a specific peptide binding site which has been identified in a crystallographic study of the T26E mutant of T4L with the covalently bound disaccharide-tetrapeptide GlcNAc-(β1→4)-MurNAc-L-Ala-D-Glu-γ-DAP-D-Ala (**V**, Figure 1; DAP: diaminopimelic acid);[6] the disaccharide units are placed in subsites C and D, while the tetrapeptide is situated between two groups of α helices of the C-terminal upper lobe of T26E formed, roughly, by residues 105−115 and 132−141, which comprise the peptide binding site. **V**

* Author to whom correspondence should be addressed. Tel. +386-61-176-0200; Fax +386-61-125-9244. E-mail primus@kihp7.ki.si.

† Keywords: muramyl peptides; lysozyme; computer docking; electrostatic potential. E-mail primus@kihp7.ki.si.

⊗ Abstract published in *Advance ACS Abstracts,* November 1, 1997.

S0095-2338(97)00009-7 CCC: $14.00

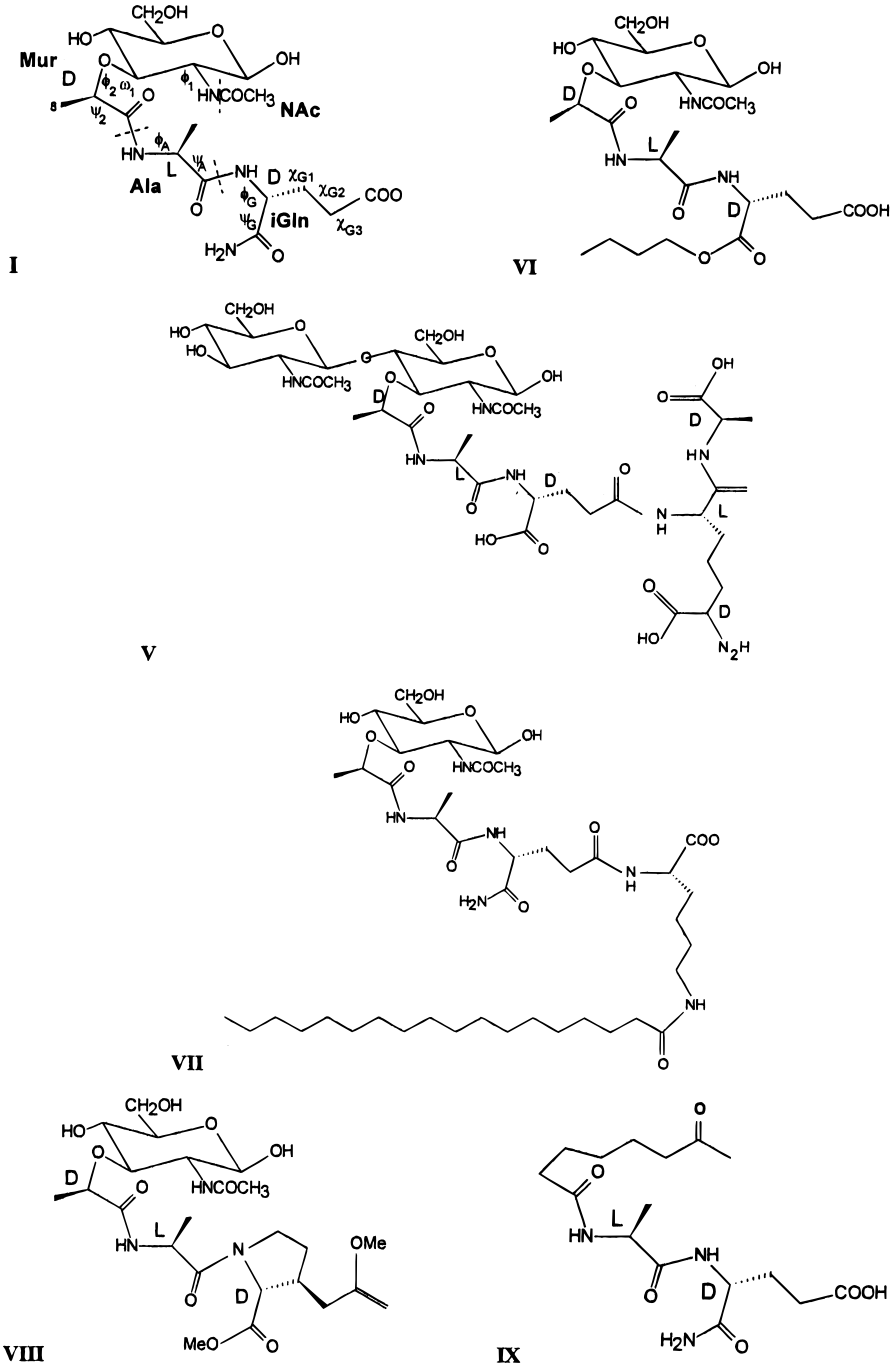

Figure 1. Chemical structures of MDP (**I**), GlnNAc-β1→4-MurNAc-L-Ala-D-iGln-γ-DAP-D-Ala (**V**), murabutide (**VI**), muroctasin (**VII**), *N*-acetylmuramyl-L-alanyl-3-(carbomethoxymethyl)-D-proline methyl ester (**VIII**), and LK409 (**IX**).

exhibits adjuvant activity and was used to derive the smaller MDP fragment **I**.[13]

It is this specific peptide binding site (PBS) that we focused on in our work. The two basic assumptions are (i) the PBS specifically binds the L-Ala-D-Glu peptide sequence that is a constituent of bacterial cell walls and (ii) the putative receptor retained or developed a similar binding mode for discriminating the same sequence. We take the PBS as a partial model for the true binding site and use it for docking of active and inactive analogues. A thorough analysis of their binding modes (i.e., interactions with the PBS' residues and patterns of intermolecular H-bonds) reveals common features of active analogues that are not entirely shared by the inactive analogues. These common features define the specific binding mode, and the resulting picture consistent with SAR data in retrospect vindicates the approach. Additional support for this undertaking is offered by published data[14] showing that exactly the same ligand **V** can be bound in nearly the same conformation by goose egg white lysozyme (which shows low sequence homology relative to T4L[15]) and the recently described transglycosilase SLT-70 of *E. coli*[14] which has a completely different overall structure of the rest of the protein.

In this paper we report the results of docking of a representative series of MFDPs, active in immunopotentiation tests, and some inactive ones, to the PBS of T26E, including (i) the diastereomers of MDP (**I**−**IV**), (ii) the original ligand **V**, (iii) murabutide[16,17] (**VI**, Figure 1), an active analogue with esterified C-terminus, (iv) muroctasin (N^2-[(*N*-acetylmuramoyl)-L-alanyl-D-isoglutaminyl]-N^6-stearoyl-L-lysine, **VII**,

Figure 1), a potent analogue with increased lipophilicity,[18] (v) the inactive rigidified analogue *N*-acetylmuramyl-L-alanyl-3-(carbomethoxymethyl)-D-proline methyl ester[19] (**VIII**, Figure 1); **VI**−**VIII** were also used in the previous paper with different numbering), and (vi) the active desmuramyl analogue LK409[20] (*N*-(7-oxooctanoyl)-L-Ala-D-iGln, **IX**, Figure 1). The interactions of the docked conformations with the PBS were analyzed by partitioning into ligand−protein group−group interactions; intermolecular hydrogen bond patterns were determined from the geometry of the complexes. The results of the present work identify the mode of binding to the PBS that leads to specificity relative to the L-Ala-D-Glu sequence, and other common features of the active analogues are observed. We propose the docked conformation of MDP as the bioactive one since the different binding modes of active and inactive MFDPs to the PBS of T26E correlate well with the corresponding structure-activity relations and, most importantly, the docked conformation of **I** is found most similar to the most probable candidate obtained by the CCLUES method (see previous paper). The results of the docking experiments confirm most of the preselected target atom groups used in the previous paper as key binding groups for recognition by the model receptor. Moreover, the striking similarity of the conformations resulting from the application of the CCLUES procedure and one resulting from the docking experiments encouraged us to propose a partial 3D pharmacophore of MDP-type immunostimulants. Since the key binding groups are definitely polar we propose the pharmacophore in the form of molecular electrostatic potential maps.

METHODS

(a) The coordinates of the bacteriophage T4 lysozyme mutant T26E[6] were used for the model of the binding site (referred to as "the site"). The covalent bond between Glu[11]-OE1 of the protein and Mur-O1 of the adduct **V** was deleted and the free ligand placed back into the site; the crystal waters were omitted since none of them appears to mediate any important interaction between the protein and the ligand. A distance-dependent dielectric constant was adopted in order to account, at least in part, for the missing solvent. A 20 ps molecular dynamics (MD) simulation at 300 K involving **V** and the site was started after initial energy minimization (EM) of the complex; the backbone atoms of the site were kept fixed, but the side chain atoms were allowed to move; the purpose of this simulation was to obtain a structure with relaxed nonbond interactions while remaining as similar as possible to the crystal structure (which, to the best of our knowledge, is the only source of experimentally obtained structural data involving binding of the peptide moiety of MDP-type peptidoglycans). Forty structures were sampled and subsequently energy minimized. The INSIGHT/DISCOVER molecular modeling package[21] was used for the MD and EM simulations. The coordinates of the site and **V** that were best in interaction energy and intermolecular hydrogen bond pattern (i.e., closest to the proposed pattern in the crystal) were used in subsequent calculations involving **I**−**IV** and **VI**−**IX**; **V** is larger in size than the latter and was used as frame for their construction and initial intermolecular alignment. The complexes of the site and manually docked conformations of **I**−**IV** and **VI**−**IX** were energy minimized and submitted to a 20 ps MD simulation at 300 K with all the atoms of the site kept fixed; in the case of

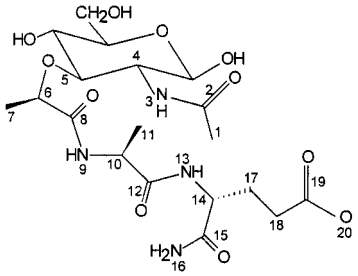

Figure 2. The division of **I** in ligand atom groups.

VII the stearoyl aliphatic chain that exits the site and enters the solvent with the third CH_2-group was replaced by a butyl group. Forty conformations of each complex were sampled and energy minimized; the complex with the most favorable intermolecular contacts and interaction energy was chosen to represent the analogue; it is referred to as "the representative complex".

For the analysis **I** was divided into 20 (=*NrLAG*) groups of atoms (referred to as ligand atom groups) which were equipped with serial numbers starting from the NAc-CH_3 and progressing toward the peptide C-terminal up to the end of the Glu side chain (Figure 2); some of the muramyl atoms were ignored in the division taking into account that the modeled binding site, even if related to T4L, most probably does not contain a glycolytic active site; moreover, the complete muramyl group is not essential for activity.[1,20,22] The same division was consistently performed with **II**−**IX**; for ligands possessing different chemical groups than **I** the same serial number for the group of atoms situated at the same position in the covalent chain was used. The DISCOVER charge groups of the amino acid residues were used for division of the site in 812 (=*NrSAG*) site atom groups.

The atom groups were used for partitioning of the intermolecular interaction energies of each representative complex in ligand-to-site group−group interaction energies which were stored in matrices $\mathbf{M_Y}$ (Y = **I**..**IX**) with dimensions *NrLAG* × *NrSAG* containing the interaction energy of ligand atom group *i* with site atom group *j* in element (*i*,*j*). The largest values of the matrix elements $\mathbf{M_Y}(i,j)$ of all active analogues (**Y** = **I**, **V**−**VII**, **IX**) were extracted for each pair (*i*,*j*) ($1 \le i \le NrLAG$, $1 \le j \le NrSAG$) to the matrix $\mathrm{Max_{Active}}$; each entry in the latter contains the most unfavorable interaction of the *i*th atom group of any active ligand with the *j*th site atom group. The matrix $\mathrm{Max_{Active}}$ was searched in each row, i.e., for each of the ligand atom groups, for the smallest value and stored in the vector (of dimension *NrLAG*) $\mathrm{Min\_Max_{Active}}$; the corresponding column numbers, i.e., the serial numbers of the site atom groups which gave the value stored in $\mathrm{Min\_Max_{Active}}$, were written to the vector $\mathrm{Index\_Min\_Max_{Active}}$. The *i*th entries in these two vectors describe how strongly and which site atom group *least unfavorably* interacts with the ligand atom group *i* in all active ligands. A negative value of $\mathrm{Min\_Max_{Active}}(i)$ means that the *i*th atom group of any active ligand has a favorable interaction energy of at least $\mathrm{Min\_Max_{Active}}(i)$ with site atom group *z* ($z = \mathrm{Index\_Min\_Max_{Active}}(i)$); a large negative value (e.g., less than −1.5 kcal/mol) denotes an important contact of all active analogues with the corresponding site atom group. A comparison with the values of $\mathbf{M_Y}(i,z)$ of the inactive analogues (**Y** = **II**−**IV**, **VIII**) may give clues to differences between active and inactive analogues at the level of

Table 1. (a) Total Interaction Energies (in kcal/mol) of Ligands **I**−**IX** with the Site and (b) the Interaction Energies (in kcal/mol) of Ligand Atom Groups (First Column) with the Site Atom Groups from the Vector Index_Min_Max$_{Active}$ (Second Column)[a]

ligand		**MMActive**	**I**	**II**	**III**	**IV**	**V**	**VI**	**VII**	**VIII**	**IX**
activity			A	I	I	I	A	A	A	I	A
					Part a. Interaction Energy						
			−108.6	−109.4	−108.5	−107.7	−230.0	−113.5	−153.9	−106.9	−66.7
					Part b						
ligand atom group	representat of site atom group										
1	CG_ASP_20	−0.58	−0.68	−0.69	−0.69	−0.69	−0.69	−0.67	−0.68	1.53	−0.58
2	NE2_GLN_105	−1.71	−1.88	−1.90	−1.91	−1.92	−1.93	−1.85	−1.87	0.82	−1.71
3	NE2_GLN_105	−0.28	−0.98	−0.96	−0.91	−0.92	−0.94	−1.01	−0.99	0.15	−0.28
4	C_GLN_105	−0.37	−1.65	−1.66	−1.69	−1.73	−1.37	−1.62	−1.64	−0.55	−0.37
5	CD_GLU_26	−0.65	−0.74	−0.76	−0.73	−0.66	−0.74	−0.69	−0.71	1.70	−0.65
6	CD2_LEU_32	−0.07	−0.08	−0.07	−0.06	−0.08	−0.09	−0.07	−0.08	−0.22	−0.13
7	NZ_LYS_35	−0.31	−0.31	−0.32	−0.31	−0.32	−0.32	−0.31	−0.31	0.30	−0.59
8	C_GLN_105	−0.82	−0.86	−0.86	−1.01	−1.01	−0.82	−0.88	−0.85	3.70	−1.07
9	CG_PRO_143	−1.63	−1.84	−1.85	−1.88	−1.86	−1.83	−1.82	−1.86	3.54	−.163
10	N_PRO_143	−1.86	−2.04	−2.14	−2.12	−2.07	−2.02	−2.02	−2.05	5.74	−1.86
11	CD_GLN_141	−0.46	−0.51	−0.47	−0.02	0.09	−0.54	−0.51	−0.53	1.24	−0.46
12	N_MET_106	−1.11	−1.11	−0.02	−0.07	−1.07	−1.14	−1.14	−1.19	−0.03	−1.36
13	SD_MET_106	−1.58	−1.77	−1.40	−1.50	−1.80	−1.73	−1.79	−1.58	2.63	−1.77
14	CD_ARG_137	−1.93	−2.23	−2.88	−3.16	−2.15	−2.14	−2.22	−1.93	3.10	−2.16
15	**NH2_ARG_137**	−0.91	−1.57	0.37	0.38	−1.39	−1.58	−1.51	−0.91	0.09	−1.51
16	C_ARG_137	−0.14	−0.19	−0.15	−0.18	−0.20	−0.29	−0.31	−0.14	0.94	−0.19
17	CD_ARG_137	−0.88	−0.98	−0.08	−0.02	−0.89	−1.00	−0.96	−1.07	−3.85	−0.88
18	C_ARG_137	−0.39	−0.53	−0.39	−0.36	−0.53	−0.51	−0.52	−0.39	−0.96	−0.51
19	**N_TRP_138**	−1.96	−2.03	−1.96	−1.21	−2.10	−2.21	−1.98	−1.96	−0.91	−2.12
20	SD_MET_106	−0.35	−0.36	−0.35	−0.37	−0.38	−1.28	−0.36	−1.45	−1.44	−0.35

[a] I.e., those site atom groups that have the least unfavorable interaction energies for all active ligands; the latter interaction energies are found in the vector Min_MaxActive which is shown in the third column. Activity denotes activity in immunopotentiation tests (A: active, I: inactive). In bold: most important contacts. Underlined: intermolecular H-bond between the groups in the first and second column.

interaction energies with the site. The group−group interaction energies and the analysis using the matrices $\mathbf{M_Y}$ (Y = **I**..**IX**) were calculated using a home written program that includes an implementation of the GROMOS87[23] energy function.

The conformation of **I** from the representative complex was also taken for calculations of the *ab initio* electrostatic potentials at the 3-21G level *in vacuo* and using the isodensity polarized continuum model (IPCM[24]) with a dielectric constant value of 5 simulating the protein medium. For these calculations the atoms O1′, O2′, O5′, C5′, C6′, and O6′ (and the corresponding hydrogen atoms) of the sugar ring (which is not an essential part of the proposed pharmacophore) were deleted and the C1′ and C2′ atoms converted to methyl groups. The *N*-acetyl group was not altered in any way; its carbonyl group is present in all ligands **I**−**VIII** and thus appears to be an essential part of the proposed pharmacophore (for the special case of **IX** see Discussion). This procedure not only is more economical but also prevents the highly polar hydroxyl groups of the sugar ring from influencing the electrostatic potential of the whole structure by different orientations. The same calculations following the same reductions of the atom set in the muramyl moieties were performed with the conformations of **II**−**IV** from their respective representative complexes. The Gaussian94[25] program was employed for the *ab initio* calculations.

(b) The conformation of **I** from the representative complex was finally compared by target atom least-RMS superpositions to the large set of conformation clusters of **I**[3] obtained with the conformational search; the structure from the latter set with the smallest RMS superposition value was sought.

RESULTS

Docking to the T26E Binding Site. In the docking simulations of the site-ligand complexes the muramyl moieties of **I**−**VIII** and the aliphatic chain of **IX** occupy the saccharide binding subsite D of T26E.[6] In the following the residues of the site are denoted with the corresponding residue numbers in superscript.

The interaction energies of the site and the ligands are shown in Table 1a; they are complex stabilizing in all cases. The values increase with the size of the ligand and can be directly compared only with the four diastereomers of MDP **I**−**IV** where nearly identical numbers are observed.

A more distinctive picture can be obtained with the group−group interaction energy analysis (see Methods) which is partly shown in Table 1b. For each ligand atom group *i* (first column) the first atom of the site atom group *z* (*z* = Index_Min_Max$_{Active}$(*i*), second column) is given; the pair (*i*,*z*) possesses the least unfavorable interaction energy for all active ligands which is stored in the vector Min_Max$_{Active}$ shown in the third column. The values vary from −1.96 to −0.14 kcal/mol; they are all negative indicating that for each ligand group at least one site atom group can be found that interacts favorably in all active ligands. In the following columns the corresponding values of $\mathbf{M_Y}(i,z)$ (**Y** = **I**..**IX**) are given. Some highly negative values and the atom types in the ligand and site atom groups (proton donors and acceptors) indicate intermolecular hydrogen bonds that are also found in the representative complexes using geometrical criteria (d(O···H−X) < 2.3 Å, θ(O,H,X) > 120°; X = N, O); such cases are marked by underlined values of $\mathbf{M_Y}(i,z)$ in Table 1b. In the series of the stereochemically related active analogues **I** and **V**−**VII** the pattern of intermolecular H-bonds is very similar to that

proposed for T26E and its (covalently bound) adduct **V**;[6] a series of intermolecular H-bonds between the ligands and the site can be identified from Table 1b engaging the NAc-CO, the iGln-CO, and the iGln-δCO groups of the ligand. Additionally, the H-bond iGln-N--OE1-Gln^{141} is observed in the representative complexes of **I**, **IV**−**VII**, and **IX**, i.e., ligands possessing D-iGln or D-Glu. The DAP and Lys side chains in **V** and **VII**, respectively, are firmly placed in the subsite that binds the diaminopimelic acid of the original substrate.[6] The terminal D-Ala of **V** and the stearoyl aliphatic chain of **VII** are nearly completely exposed to the solvent.

The binding of the active desmuramyl analogue **IX** in its representative complex closely corresponds to **I** in the Ala and iGln moiety; in the aliphatic chain at the N-terminal the correspondence is naturally low except in the case of the 7-oxo carbonyl end of the chain that engages the same NE2-Gln^{105} group in intermolecular H-bonding as NAc-CO in **I**−**VII**. In contrast, the rigidified inactive analogue **VIII** weakly corresponds to the binding modes of **I** to the site even in the Mur moiety which is common to both. The discrepancy is not caused by the esterified C-terminal; the latter presents no obstacle in binding to the site since the butyl chain is oriented into the solvent; the same is also exemplified by **VI**. The main cause is the bulky Pro side chain overlapping with the site residues Trp^{138} and Gln^{141} and forcing the iGln residue of **VIII** away from the subsite that is occupied by the iGln residue of **I**.

Special focus was put on the binding modes of the series of MDP diastereomers **I**−**IV**. Significant differences occur upon changing the iGln configuration from D (**I**, **IV**) to L (**II**, **III**); most notably, the position of the iGln side chain is retained, while the H-bonds iGln-CO--H2NE-Arg^{137} and iGln-N--OE1-Gln^{141} are broken and the interactions between the corresponding groups of atoms are turned into a weak repulsion. The side chain of Gln^{141} shifts toward the N-terminal of the ligands and is found engaged in the H-bond Ala-CO--OE2-Gln^{141}. The effects of the change of L-Ala → D-Ala in analogues **III** and **IV** are less clear. In the representative complexes of **III** and **IV** the H-bonds Ala-N--OC-Gln^{105} are present; otherwise only a slight weakening of the interactions Ala-CH_3-Gln^{141} side chain is evident from Table 1b. A closer look at the $\mathbf{M_{III}}$ and $\mathbf{M_{IV}}$ matrices also showed increased repulsion of Ala-CαH with the side chain of Trp^{138}.

Comparison of the Candidates for the Bioactive Conformation (Previous Paper) and the Docked Conformation of I. When compared to the 1035 conformation clusters of **I**[3] by target atom least-RMS superposition the docked conformation from the representative complex of **I** was found most similar (i.e., having the lowest RMS superposition value of 0.8 Å) to the representative of cluster no. 825 (Figure 3). The latter is also found in the subset of the three most probable candidates for the bioactive conformation; it was selected as the final most probable candidate by the refined CCLUES procedure (see previous paper).

A comparison of selected dihedral angles is given in Table 2; some significantly different values are found (e.g., φ_{Ala}, φ_{iGln}, $\chi_{1\mathrm{iGln}}$, and $\chi_{3\mathrm{iGln}}$). However, this is not surprising since it is the position of target atoms in space (which is indeed similar) that is compared and not the whole conformation (Figure 3).

The question also arises as to why exactly the same conformation as docked to the model receptor is not present in the set of conformation clusters of **I**. In this context, the energy of the docked conformation of **I** *in vacuo* becomes interesting; it computes to a high energy of 32.6 kcal/mol (relative to the energy of the representative of cluster no. 1) and lowers its energy to 13.02 kcal/mol upon energy minimization (but significantly changing its conformation in the process: RMS on all atoms 2.16 Å; the final conformation is not similar to the representative of cluster no. 825). The docked conformation (like the representative of cluster no. 825) is therefore a high-energy structure when isolated and is part of the conformational space that is not extensively sampled by the conformational search procedure used.[3] This result is also in agreement with the previous paper and the results of a systematic study by Nicklaus *et al.*[26] which reports that the bound conformation

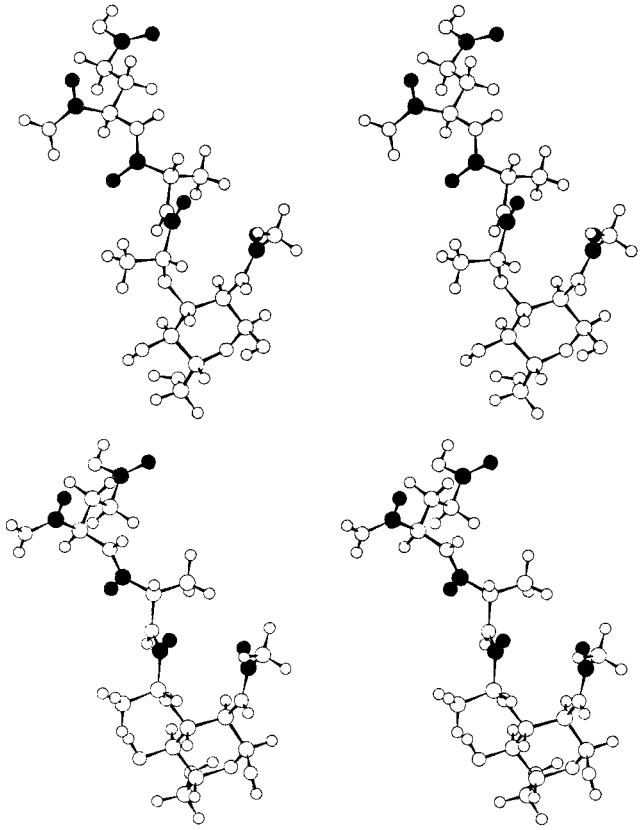

Figure 3. Comparison of the docked conformation of **I** from its representative complex (top) and the final candidate for the bioactive conformation obtained by the CCLUES procedure in the previous paper[2] (bottom). The target atoms are marked black.

Table 2. Values of Some Dihedral Angles (in deg) of the Docked Conformation of **I** from Its Representative Complex and of the Final Candidate for the Bioactive Conformation Obtained by the CCLUES Procedure[2]

dihedral angle	docked	CCLUES
NAc-C′- -Mur-N- -Mur-C2- -Mur-C3	−1.37	−109
Mur-N- -Mur-C2- -Mur-C3- -Mur-C9	78	65
Mur-C2- -Mur-C3- -Mur-O3- -Mur-C9	−84	−99
Mur-C3- -Mur-O3- -Mur-C9- -Mur-C10	111	137
Mur-O3- -Mur-C9- -Mur-C10- -Ala-N	32	34
φ_{Ala}	167	−140
ψ_{Ala}	170	142
φ_{iGln}	83	139
ψ_{iGln}	−109	−112
$\chi_{1\mathrm{iGln}}$	167	72
$\chi_{2\mathrm{iGln}}$	−170	178
$\chi_{3\mathrm{iGln}}$	−166	−89
$\chi_{4\mathrm{iGln}}$	−176	−179

of a flexible ligand is not usually any of the low-energy minima found *in vacuo* and that the protein-bound conformation may be found virtually anywhere on the energy surface.

Recently a conformational analysis in water using NMR, simulated annealing, and molecular dynamics of a disaccharide-pentapeptide closely related to **V** has been published;[27] the torsional angles of the peptide chain from neither the simulated annealing nor the MD simulations coincide with those found in the docked conformation of **V**. This is again an indication that protein-bound conformations of ligands are not necessarily found in the set of low-energy structures proposed in solution.

Proposal of an MDP Pharmacophore. The results summarized above show impressive correlations between the bioactivity of **I**−**IX** and their modes of binding to the T26E peptide binding site. We propose that the PBS of T26E may be used as a first model of the putative MFDPs' binding protein and the docked conformation of MDP (**I**) as the bioactive conformation as far as the dipeptide core is concerned. From our model a series of pivotal contacts of **I** with the "receptor" can be derived. The hydrogen bonding of 7-oxo of **IX** to H2NE-Gln^{105} and its bioactivity offers a simple explanation in that the contact with H2NE-Gln^{105} is essential for triggering the response; all other active analogues fulfill this requirement using their NAc-CO group. Other pivotal contacts, namely of iGln-CO, iGln-N, and iGln-δCO with Arg^{137}, Gln^{141}, and Trp^{138}, respectively, are indicated by the results of the series **I**−**VII** (Table 1b).

The geometry of the docked conformation of **I** and the resulting spatial positions of the above contacting groups, namely NAc-CO, iGln-CO, iGln-N, and iGln-δCO, are the main features of the MDP pharmacophore that we propose. Its core is formed by the L-Ala-D-iGln dipeptide part; from the muramyl moiety only the NAc group enters the pharmacophore while the rest is obviously not essential.[1,20] Owing to the polar nature of the pivotal contacts an electrostatic description of the pharmacophore is most suitable and is shown in Figure 4a, with the electrostatic potentials calculated at the HF/3-21G (which has been used for dipeptides previously[28]) level with the IPCM method[24] using the dielectric constant value of 5; the muramyl moiety (except for the NAc group) has been omitted from the calculations (v. infra). The four pivotal contacting groups are sites of large gradients in the potential. For comparison, the electrostatic potential of the inactive diastereomer **II** is shown in Figure 4b; the differences are most obvious at the Cα-atoms of iGln where the inversion of chirality and resulting rearrangement of the C-terminal carboxamide group occurs; common features include the *N*-acetyl groups in the muramyl moiety and the side chain carboxyl groups of iGln.

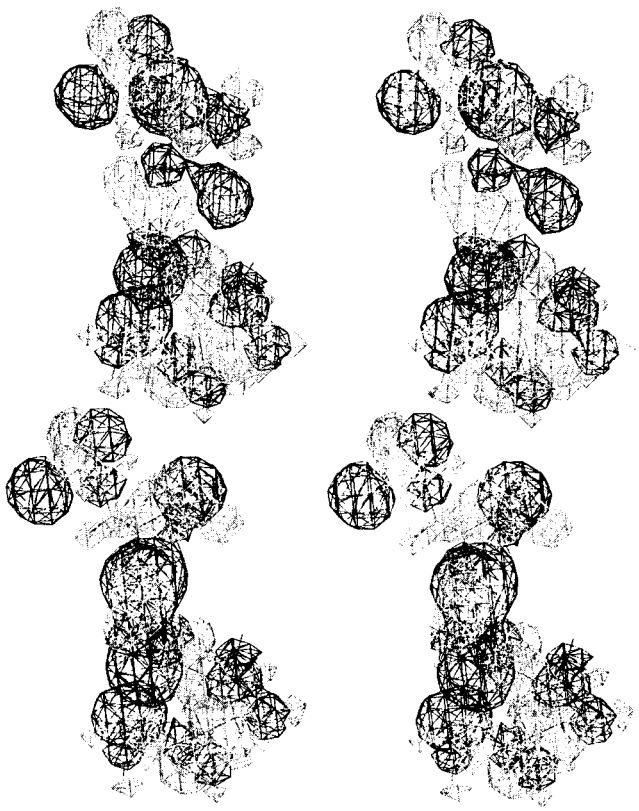

Figure 4. The isopotential contours of the electrostatic potential of the modified (see Methods) molecules of **I** (a, top) and **II** (b, bottom) calculated at the HF/3-21G level with the IPCM method[24] using the dielectric constant value of 5 (dark: isopotential contours of −35 kcal/mol/As; light: isopotential contours of 35 kcal/mol/As).

DISCUSSION

The results of the computer docking of MFDPs **I**−**IX** to the T26E binding site, summarized in Tables 1 and 2, support the viability of the hypothesis suggesting that similar molecular determinants of recognition are involved in eliciting cell response and in activating the v-type lysozymes. The values of the interaction energies (Table 1a) are not significant in this respect since the solvent is only approximately taken into account ($\epsilon = r$) and the method does not necessarily lead to global minima of the complexes; furthermore, time-consuming calculation methods that allow an estimate of the free energies of binding[29] are not justified in this work since no experimental data for comparison are available. What matters are the free energies of transfer from aqueous solution to the protein binding site. In the case of receptor activation, the stability of the ligand−receptor complex as expressed by negative ΔG does not differentiate between receptor activating (agonistic) and binding but not activating (antagonistic) ligands; it is the proper mode of binding that is required for activation.

The molecular dynamics of the docked ligands **I**−**IV** and **VI**−**IX** has been carried out with the active site atoms fixed; a moving active site introduces new degrees of freedom which make the common binding motifs of the active ligands much more difficult to trace without necessarily introducing additional insights. The peptide-binding site in this study serves as a model for the binding site of the putative receptor; allowing it to adapt to each of the flexible ligands would impress onto the model too much of the characteristics of the T26E lysozyme. For the same reason a solvated environment in the simulations was omitted.

The configurational differences of the MDP analogues are indeed reflected in the modes of binding that can be identified by analysis of the ligand-to-site group−group interactions. Most importantly, the site favorably but differently binds ligands with D- and L-configurations of Glu (and Ala, to a lesser extent) which is also the main factor deciding on biological activity. That the various diastereomers of MDP **II**−**IV** do even antagonize the activation by **I**[1,30,31] is in agreement with the presently demonstrated possibility of

binding for each of them. The motif including the intermolecular H-bonds iGln-CO--H2NE--Arg[137], iGln-N--OE1--Gln[141], and iGln-δCO--HN-Trp[138] is probably the main feature of selectivity toward the D-configuration of iGln and is displayed by all active analogues. The selectivity of the T4L peptide binding site could be verified experimentally by measuring the digestion rate of v-type lysozymes upon substrates containing, e.g., L-Ala-L-Glu dipeptide sequences attached to the muramic acid; a slower digestion rate would be expressed. To the best of our knowledge, such a study has not been reported in the literature. In absence of experimental data a free energy perturbation study containing the conversion D-Glu → L-Glu is underway in our laboratory and will be reported at a later date.

The last of the above listed H-bonds, iGln-δCO--HN-Trp[138], is displayed by all analogues (except of **VIII** which does not fit the PBS due to steric overlap); its role as pivotal contact is suggested by the inactivity of the MFDPs that have a carbonyl function missing at the third carbon of the side chain (e.g., MurNAc-L-Ala-D-Asp).

The significance of the muramyl moiety for binding and activity also deserves attention. Lysozymes have evolved for the hydrolysis of β1→4-glycosidic linkages in oligosaccharides and/or transglycosylation.[5] This and the structure of **I** suggest that the muramyl moiety should be important for binding to the putative receptor and subsequent activation. However, the immunostimulant activities of several desmuramyl MFDPs witness to the contrary.[20,22] In this context the importance of the 2-acetamido group of **I** or an equivalent one in other MFDPs for binding to the lysozymes and for the activity of the MFDPs has to be underscored. Recently, a series of desmuramyl analogues containing an aliphatic chain with a carbonyl group have been synthesized[20] which shows that a carbonyl group five bonds apart from the Ala-NH is necessary for activity, at least in this dipeptide series. In agreement with these experimental findings the present docking simulations demonstrate that efficient binding occurs even in the absence of the muramyl moiety; in the case of **IX** the 7-oxo carbonyl engages the same H2NH-Gln[105] ϵN as the 2-acetamido group of the other analogues in the series that possess the muramyl moiety. The need for the 7-oxo in the LK409 series somehow contrasts the results obtained with N^2-[*N*-(*N*-lauroyl-L-alanyl)-γ-D-glutamyl]-N^6-(glycyl)-DD,LL-2,6-diaminopimelamic acid[22] and related desmuramyl compounds that are lacking that group. However, these compounds are tetrapeptides, and it is possible that they bind with additional modes that compensate for the loss of the former contact by contacts of the C-terminal part to the binding site.

In the present binding scheme we have throughout considered neutral H-bonding; however, the contact of the C-terminal Glu carboxyl group of **V** with H2NE-Arg[137] is most likely to be ionic and should contribute more binding energy. The other MFDPs considered have either an carboxamidic (**I**−**IV**, **VII**, **IX**) or esterified (**VI** and **VIII**) C-terminal. However, recent work of Merhi et al.[32] suggests that these groups are hydrolyzed in the macrophages.

CONCLUDING REMARKS

Owing to the simple simulation models presently used the pivotal contacts and the partial pharmacophore remain for awhile a working hypothesis. It can be verified only by structural and mechanistic elucidation of the immunostimulation by MFDPs; the results of Nesmeyanov *et al.*[33] are promising to evolve in this direction. Although the docking experiments to the T4 lysozyme binding site produce some promising results, the hypothesis of structure relations between this site and that of the putative receptor have to be considered with caution. On the positive side is the discrimination between the diastereomeric dipeptide core and between the active and inactivde analogues. On the negative side is the fact that selected analogues from the series **I**−**IX** do not cocrystallize with the T4 lysozyme (Matthews, B. W.; Weaver, L. H. Personal communication) and that covalent binding to the T26E mutant was necessary to obtain crystals with **V** bound to subsite D. This fact suggests that the few hydrogen bonds identified in the present docking will not suffice for a ΔG as required for at least μm binding constants.

It is nevertheless worth stressing that even this initial, rough approach to binding leads to a consistent picture that explains most of the known SAR data. One of the impressive findings of the docking experiments is the impossibility of placing **VIII** in the T26E peptide binding subsite in a mode corresponding to **I** and **V**−**VII**; this result nicely correlates with the inactivity of **VIII** in immunopotentiating tests.[19] The most reassuring finding of the present docking simulations is the convergence of its results with the results of the completely independent CCLUES method (previous paper); we consider this as a strong support for the lysozyme PBS using a conformation unique to the L-Ala-D-iGln analogue of the MDP series in presenting the target atoms to achieve specific binding. This poses another argument in favor of the suitability of the T26E peptide binding site as a model for the putative receptor since the latter will most probably also use the conformation essentially accessible only to the L,D-diastereomer of **I** for specific binding and signal triggering. We emphasize that the proposed pharmacophore is only partial in that it covers the peptide core of MDP. The existing structure−activity relations strongly suggest that, in addition to the dipeptide core, substituents at either the Ala or Glu sides are required for activity.[1] However, the structural variety of these substituents indicates that they rather contribute in a nonspecific manner to ΔG while specificity is entailed primarily by the dipeptide core as shown by the docking experiments.

We also emphasize that three of the five target atom groups in the previous paper turn out to be pivotal contact groups in the present model. Further discussion of structural and evolutionary relations between the putative receptor and v-type lysozymes will be published elsewhere. Clearly, direct support for the hypothetical structural relations can be provided only by the structure elucidation of the MDP receptor.

ACKNOWLEDGMENT

This work was supported by the Ministry of Science and Technology of Slovenia and (in part) by the Lek Works, Ljubljana. We are obliged to Prof. Brian W. Matthews and Dr. Larry H. Weaver (University of Oregon, Eugene) for the coordinates of T26E and for communicating the results of attempted cocrystallization. We thank Prof. Vladimir Nesmeyanov (Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Moscow) for the preprint of ref 33, Prof.

Roger Paine for stimulating discussions, and Dr. Janez Mavri for help with the *ab initio* calculations. We wish to thank both anonymous reviewers for valuable suggestions.

REFERENCES AND NOTES

(1) Adam, A.; Lederer, E. Muramyl peptides: immunomodulators, sleep factors, and vitamins. *Med. Res. Rev.* **1984**, *4*, 111−152.
(2) Pristovšek, P.; Kidrič, J.; Hadži, D. Proposal of a 3D Peptide Pharmacophore of Muramyl Dipeptide-Type Immunmostimulants. 1. Conformational Search of Active and Inactive Analogues. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 971−976.
(3) Pristovšek, P.; Kidrič, J.; Hadži, D. Bioactive conformations of small peptides: a method for selection of candidates based on conformations of active and inactive analogues and its application to muramyl dipeptide. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 633−639.
(4) Anderson, W. F.; Gruetter, M. G.; Remington, S. J.; Weaver, L. H.; Matthews, B. W. Crystallographic determination of the mode of binding of oligosaccharides to T4 bacteriophage lysozyme: implications for the mechanism of catalysis. *J. Mol. Biol.* **1981**, *147*, 523−543.
(5) Jolles, P.; Jolles, J. What's new in lysozyme research? *Mol. Cell. Biochem.* **1984**, *63*, 165−189.
(6) Kuroki, R.; Weaver, L. H.; Matthews, B. W. A covalent enzyme-substrate intermediate with saccharide distortion in a mutant T4 lysozyme. *Science* **1993**, *262*, 2030−2033.
(7) Cheetham, J. C.; Artymiuk, P. J.; Phillips, D. C. Refinement of an enzyme complex with inhibitor bound at partial occupancy − hen egg-white lysozyme and tri-N-acetylchitotriose at 1.75 Å resolution. *J. Mol. Biol.* **1992**, *224*, 613−628.
(8) Hadfield, A. T.; Harvey, D. J.; Archer, D. B.; MacKenzie, D. A.; Jeenes, D. J.; Radford, S. E.; Lowe, G.; Dobson, C. M.; Johnson, L. N. Crystal structure of the mutant D52S hen egg white lysozyme with an oligosaccharide product. *J. Mol. Biol.* **1994**, *243*, 856−872.
(9) Song, H.; Inaka, K.; Maenaka, K.; Matsushima, M. Structural changes of active site cleft and different saccharide binding modes in human lysozyme co-crystallized with hexa-N-acetyl-chitohexaose at pH 4.0. *J. Mol. Biol.* **1994**, *244*, 522−540.
(10) Warshel, A.; Levitt, M. Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilisation of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* **1974**, *103*, 227−249.
(11) Levitt, M. On the nature of the binding of hexa-N-acetylglucosamine subsarte to lysozyme. In *Peptides, Polypeptides and Proteins*; Blout, E. R., Bovey, F. A., Goodman, M., Lotan, N., Eds.; Wiley & Sons: New York, 1996; pp 99−113.
(12) Jensen, H. B.; Kleppe, G.; Schindler, M.; Mirelman, D. The specificity requirements of bacteriophage T4 lysozyme. *Eur. J. Biochem.* **1976**, *66*, 319−325.
(13) Ellouz, F.; Adam, A.; Ciorbaru, R.; Lederer, E. Minimal structural requirements for adjuvant activity of bacterial peptidoglycan derivatives. *Biochem. Biophys. Res. Commun.* **1974**, *59*, 1317−1325.
(14) Thunnissen, A. M. W. H. The hole story -structure and function of the 70 kDa soluble lytic transglycosylase from Escherichia coli. Ph.D. Thesis, University of Groningen, NE. 1995.
(15) Gruetter, M. G.; Weaver, L. H.; Matthews, B. W. Goose lysozyme structure: an evolutionary link between hen and bacteriophage lysozymes? *Nature* **1983**, *303*, 828−831.
(16) Fermandjian, S.; Perly, B.; Level, M.; Lefrancier, P. A comparative 1H NMR study of MurNac-L-Ala-D-iGln (MDP) and its analogue murabutide: evidence for a structure involving two successive beta-turns in MDP. *Carbohydr. Res.* **1987**, *162*, 23−32.
(17) Bahr, G. M.; Darcissac, E.; Bevec, D.; Dukor, P.; Chedid, L. Immunopharmacological activities and clinical development of muramyl peptides with particular emphasis on murabutide. *Int. J. Immunopharmac.* **1995**, *17*, 117−131.
(18) Classen, H. G.; Schramm, V. Muroctasin. *Arzneim.-Forsch. /Drug Res.* **1996**, *38*, 951−1074.
(19) Boulanger, Y.; Tu, Y.; Ratovelomanana, V.; Purisima, E.; Hanessian, S. Conformation of MurNAc-L-Ala-D-iGln (MDP) and of a constrained analog using 1H NMR data and molecular modeling. *Tetrahedron* **1992**, *46*, 8855−8868.
(20) Sollner, M.; Kotnik, V.; Pečar, S.; Štalc, A.; Simčič, S.; Povšič, L.; Herzog-Wraber, B.; Klampfer, L.; Ihan, A.; Grossman, P. Apyrogenic syntetic desmuramyl dipeptide, LK409, with immunomodulatory properties. *Agents Action* **1993**, *38*, 275−280.
(21) BIOSYM/MSI; 9685 Scranton Rd, San Diego, CA 92121-3752, 1995.
(22) Migliore-Samour, D.; Bouchaudon, J.; Floc'h, F.; Zerial, A.; Ninet, L.; Werner, G. H.; Jollés, P. A short lipopeptide, representative of a new family of immunological adjuvants devoid of sugar. *Life Sci.* **1980**, *26*, 883−888.
(23) van Gunsteren, W. F.; Berendsen, H. J. C. *Groningen Molecular Simulation* (*GROMOS*) *Library Manual*; Biomos B.V. Nijenborgh 16, NL 9747 AG Groningen, 1987; pp 1−229.
(24) Wiberg, K. B.; Rablen, P. R.; Rush, D. J.; Keith, T. A. Experimental and theoretical studies of the effect of the medium on the rotational barriers for N,N-dimethylformamide and N,N-dimethylacetamide. *J. Am. Chem. Soc.* **1995**, *117*, 4261.
(25) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, W. P. M.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian 94, Revision B.3*; Gaussian, Inc.: Pittsburgh, PA, 1995.
(26) Nicklaus, M. C.; Wang, S.; Driscoll, J. S.; Milne, G. W. A. Conformational Changes of Small Molecules Binding to Proteins. *Bioorg. Med. Chem.* **1995**, *3*, 411−428.
(27) Matter, H.; Szilagyi, L.; Forgo, P.; Marinić, Ž.; Klaić, B. Structure and dynamics of a peptidoglycan monomer in aqueous solution using NMR spectroscopy and simulated annealing calculations. *J. Am. Chem. Soc.* **1997**, *119*, 2212−2223.
(28) Perczel, A.; McAllister, M. A.; Czaszar, P.; Csizmadia, I. G. Peptide models. IX. A complete conformational set of For-Ala-Ala-NH2 from ab initio computations. *Can. J. Chem.* **1994**, *72*, 2050−2070.
(29) Reynolds, C. A.; King, P. M. Free-energy perturbation methods in computer-aided molecular design. In *Computer-Aided Molecular Design*; Richards, W. G., Ed.; IBC Technical Services Ltd.: London, 1989; pp 51−60.
(30) Leclerc, C.; Chedid, L. Are there receptors for MDP at the macrophage surface? *Int. J. Immunotherapy* **1986**, *II*, 109−114.
(31) Pabst, M. J.; Johnson Jr., R. B. Increased production of superoxide anion by macrophages exposed in vitro to muramyl dipeptide. *J. Exp. Med.* **1980**, *151*, 101−114.
(32) Merhi, G.; Coleman, A. W.; Devissaguet, J.-P.; Barratt, G. M. Synthesis and immunostimulating properties of lipophilic ester and ether muramyl peptide derivatives. *J. Med. Chem.* **1995**, *39*, 4483−4488.
(33) Nesmeyanov, V. A.; Golovina, T. N.; Khaidukov, S. V.; Shebzuchov, Y. V. Muramyl peptide-binding proteins of macrophages: identification and characterization. In *Peptides in immunology*; Schneider, C. H., Ed.; Wiley & Sons: Chichester, 1996; pp 291−294.

CI970009T