J|A|C|S
ARTICLES
Published on Web 03/22/2006

Concepts in Receptor Optimization: Targeting the RGD Peptide

Wei Chen, Chia-en Chang,† and Michael K. Gilson*

Contribution from the Center for Advanced Research in Biotechnology, University of Maryland Biotechnology Institute, 9600 Gudelsky Drive, Rockville, Maryland 20850

Received September 26, 2005; E-mail: gilson@umbi.ume.edu

Abstract: Synthetic receptors have a wide range of potential applications, but it has been difficult to design low molecular weight receptors that bind ligands with high, "proteinlike" affinities. This study uses novel computational methods to understand why it is hard to design a high-affinity receptor and to explore the limits of affinity, with the bioactive peptide RGD as a model ligand. The M2 modeling method is found to yield excellent agreement with experiment for a known RGD receptor and then is used to analyze a series of receptors generated *in silico* with a *de novo* design algorithm. Forces driving binding are found to be systematically opposed by proportionate repulsions due to desolvation and entropy. In particular, strong correlations are found between Coulombic attractions and the electrostatic desolvation penalty and between the mean energy change on binding and the cost in configurational entropy. These correlations help explain why it is hard to achieve high affinity. The change in surface area upon binding is found to correlate poorly with affinity within this series. Measures of receptor efficiency are formulated that summarize how effectively a receptor uses surface area, total energy, and Coulombic energy to achieve affinity. Analysis of the computed efficiencies suggests that a low molecular weight receptor can achieve proteinlike affinity. It is also found that macrocyclization of a receptor can, unexpectedly, *increase* the entropy cost of binding because the macrocyclic structure further restricts ligand motion.

2. Introduction

Low molecular weight, synthetic receptors that bind targeted ligands with high affinity should be useful in a range of applications such as chemical detection, separation, and catalysis. There are also intriguing possibilities for biomedical applications, not only in the formulation of pharmaceuticals but also as therapeutic agents in their own right; see, e.g., refs 1−3. Synthetics offer potential advantages over proteins, arguably their chief competitors, including greater physical and chemical stability, lower molecular weight, and a far more varied selection of chemistries for the creation of structure and functionality.

Elegant receptors have been synthesized that bind ligands in aqueous and organic environments, but it has proven difficult to reach the high affinities that are desirable for many applications and that are routinely provided by proteins. (See Table 1 and Figure 25 in ref 4.) It might be argued that synthetic receptors are simply too small to achieve nanomolar dissociation constants routinely. On the other hand, it should be possible for a small, artificial receptor to wrap a ligand as thoroughly as a protein does and to form equally complementary noncovalent interactions, without carrying the bulky baggage of a protein. Indeed, there is no fundamental argument against the existence of low molecular weight receptors that will bind a ligand with affinities similar to those achieved by proteins. It is thus of interest to probe the upper limits of the size efficiency of synthetic receptors, where size efficiency may be defined as the binding free energy per heavy atom of the receptor, in analogy with the concept of ligand efficiency for small molecules that bind proteins.[5]

Despite years of broad and intense research in host−guest and supramolecular chemistry, surprisingly little effort has been devoted to techniques for the design of synthetic receptors; the HostDesigner tool for *de novo* design of small receptors appears to be the only program in its category until now.[6,7] This state of affairs may stem from the absence of any effective method of predicting the affinity of a candidate receptor for its intended ligand, as recently noted.[8] This situation is changing, however, with the development of the M2 method of computing affinities. M2 accounts not only for direct interactions between the receptor and its ligand but also for solvent and entropic contributions, providing binding free energies that are accurate to within about 1.5 kcal/mol for a range of experimentally characterized systems.[9,10] It also provides detailed information about the

† Current address: Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, CA.

(1) Adam, J. M.; Bennett, D. J.; Born, A.; Clark, J. K.; Feilden, H.; Hutchinson, E. J.; Palin, R.; Prosser, A.; Rees, D. C.; Rosair, G. M.; Stevenson, D.; Tarver, G. J.; Zhang, M.-Q. *J. Med. Chem.* **2002**, *45*, 1806−1816.
(2) Zhang, M. Q. *Drugs of the Future* **2003**, *28*, 347−354.
(3) Epemolu, O.; Bom, A.; Hope, F.; Mason, R. *Anesthesiology* **2003**, *99*, 632−637.
(4) Houk, K. N.; Leach, A. G.; Kim, S. P.; Zhang, X. Y. *Ang. Chem., Int. Ed.* **2003**, *42*, 4872−4897.
(5) Hopkins, A. L.; Groom, C. R.; Alex, A. *Drug Discov. Today* **2004**, *9*, 430−431.
(6) Hay, B. P.; Firman, T. K. *Inorg. Chem.* **2002**, *41*, 5502−5512.
(7) Bryantsev, V. S.; Hay, B. P. *J. Am. Chem. Soc.* **2006**, *128*, 2035−2042.
(8) Menger, F. M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 4818−4822.

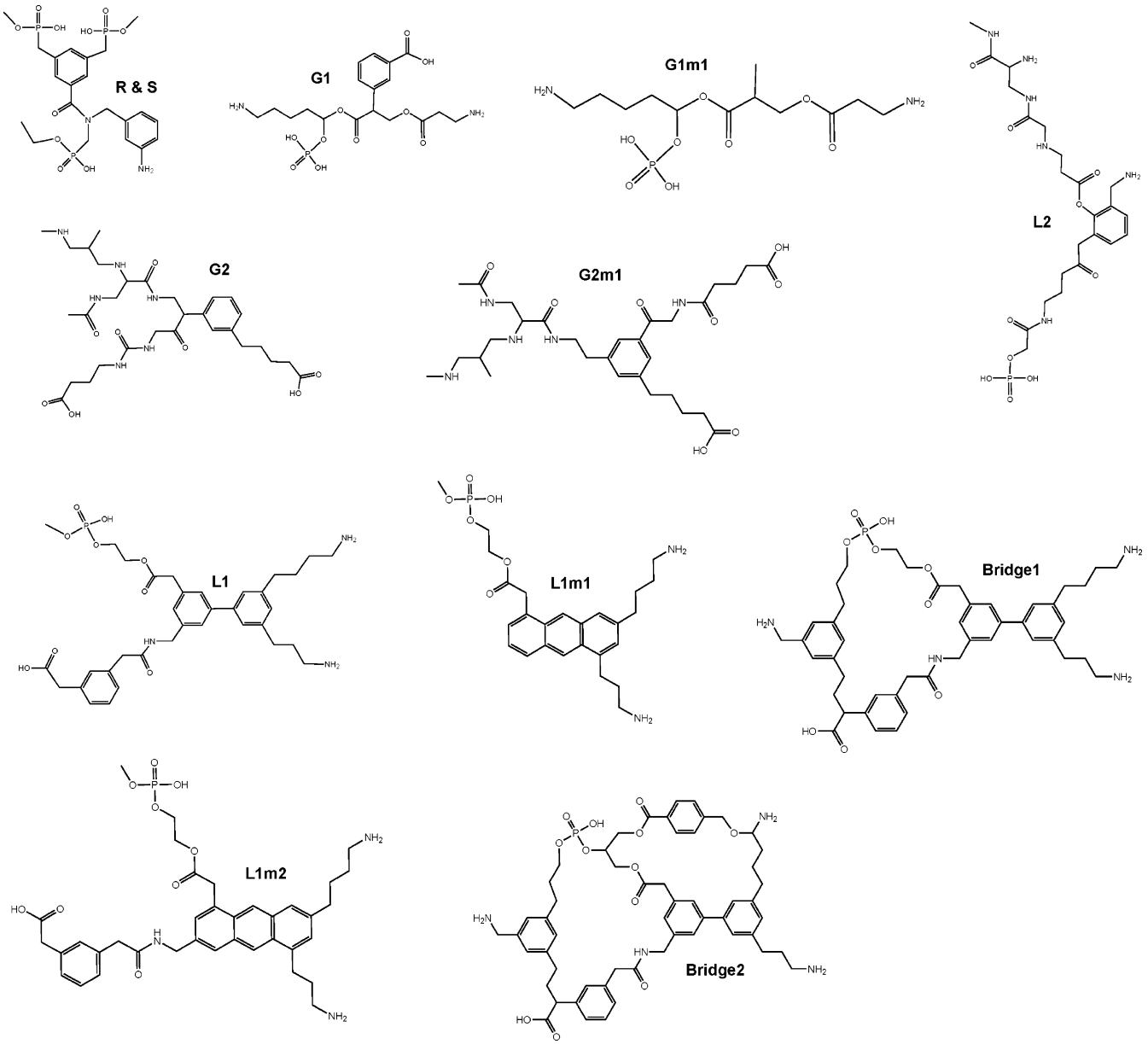

Figure 1. Chemical structures of **R&S** receptor and designed receptors.

conformational preferences of the free and bound species, and about the various energy and entropy changes that accompany binding, information which can guide further optimization of the specific receptor and which can furthermore provide insights into the nature of molecular recognition that are more broadly applicable.

The RGD peptide occurs at the surfaces of a number of extracellular proteins and confers the ability to bind integrins, cell-surface transmembrane proteins that provide for the adhesion of cells to their external matrix. (See, e.g., refs 11−13 and additional references therein.) Integrin binding furthermore generates intracellular signals influencing multiple biological functions, such as development, immunity, haemostasis, and other physiologic processes; for a concise review, see ref 14. Therefore, peptides containing the RGD sequence (see, e.g., ref 15 and references therein) and nonpeptidic analogues (see, e.g., ref 16) that also bind integrins are of great interest as steps toward new medications for a range of conditions including thrombosis, inflammation, and cancer. Recently, Rensing and Schrader have suggested using compounds that bind the RGD peptide itself and thereby block its interaction with integrins.[17] These authors created a synthetic receptor (**R&S** in Figure 1) which binds both linear RGD and a cyclic derivative of RGD, but not a synthetic RGD mimic. The receptor's affinity for RGD is substantial, −4.3 kcal/mol, but far less than might be expected for a protein.

(9) Chang, C.-E.; Gilson, M. K. *J. Am. Chem. Soc.* **2004**, *126*, 13156−13164.
(10) Chen, W.; En Chang, C.; Gilson, M. K. *Biophysical Journal* **2004**, *87*, 3035−3049.
(11) Hynes, R. O. *Cell* **1987**, *48*, 549−554.
(12) Ruoslahti, E. *Annu. Rev. Cell. Dev. Biol.* **1996**, *12*, 697−715.
(13) Arnaout, A.; Goodman, S. L.; Xiong, J.-P. *Curr. Opin. Cell Biol.* **2002**, *14*, 641−651.
(14) Hynes, R. O. *Cell* **2002**, *110*, 673−687.
(15) Haubner, R.; Gratias, R.; Deifenbach, B.; Goodman, S. L.; Jonczyk, A.; Kessler, H. *J. Am. Chem. Soc.* **1996**, *118*, 7461−7472.
(16) Stilz, H. U.; Guba, W.; Jablonka, B.; Just, M.; Klingler, O.; König, W.; Wehner, V.;; Zoller, G. *J. Med. Chem.* **2001**, *44*, 1158−1176.
(17) Rensing, S.; Schrader, T. *Org. Lett.* **2002**, *4*, 2161−2164.

The RGD sequence itself is, arguably, a suboptimal drug target, owing to the high concentration of RGD-containing proteins, such as fibrinogen. Nonetheless, targeting RGD represents an intriguing and instructive challenge in biomimetic receptor design for a highly ionic ligand. The present paper thus builds on the elegant host−guest studies of Rensing and Schrader, first documenting that the M2 method yields excellent agreement with experiment for their designed receptor, and then analyzing a series of *de novo* designed receptors for RGD. The results provide insight into receptor affinity and efficiency, helping to explain why it is difficult to design high affinity receptors and highlighting the reasons a given receptor may perform particularly poorly or particularly well.

3. Methods

3.1. Calculation of Binding Affinities and Free Energy Components. As previously described,[9] the M2 method computes binding affinities by computing the standard chemical potentials of the free receptor, the free ligand, and their complex, and taking the difference to obtain the standard free energy of binding. The standard chemical potential of each molecular species is obtained as a sum of contributions from the low-energy conformations of the species. These conformations are identified with the Tork conformational search algorithm,[18] and a symmetry-corrected method is used to ensure that no conformation is double-counted in the free energy sums.[19] The contribution of each unique energy well to the free energy is computed with an augmented form of the harmonic approximation, the Harmonic Approximation/Mode Scanning (HA/MS) method.[20] Note that the ligand, the receptor, and their complex are treated as fully flexible during these calculations.

These calculations also yield a decomposition of the binding free energy change ΔG° into the change in the Boltzmann averaged energy terms and the change in configurational entropy:[9,10]

$$\Delta G^\circ_{\mathrm{calcd}} = \Delta\langle U + W\rangle - T\Delta S^\circ_{\mathrm{config}} \quad (1)$$

Here U and W are, respectively, the potential energy and the solvation energy as a function of conformation, the angle brackets indicate Boltzmann averages, T is absolute temperature, and $\Delta S^\circ_{\mathrm{config}}$ is the change in configurational entropy upon binding. The change in configurational entropy comprises changes in the so-called rotational and translational entropy of the molecules upon binding, as well as the change in internal entropy, which is primarily associated with bond rotations. However, it does not include the change in solvent entropy and therefore cannot be compared directly with the experimentally measured binding entropy. The change in the Boltzmann averaged energy terms can be further decomposed into individual energy terms, such as the change in the mean Coulombic energy $\Delta\langle U_{\mathrm{coul}}\rangle$.

The potential energy, U, as a function of conformation is computed using the CHARMM force field[21] for bond-stretch, angle-bend, dihedral, and van der Waals terms. Coulombic energies are computed with partial atomic charges from the VC/2004[22] charging method as implemented in the program Vcharge.[23] The solvation energy, W, as a function of conformation is approximated with a generalized Born (GB) electrostatics model[24] during the Tork conformational search and the HA/MS analysis of each energy well. Then the solvation energy for each energy minimum is corrected toward a more accurate model by subtracting the generalized Born energy and substituting the result of a finite-difference solution of the linearized Poisson−Boltzmann equation, solved with the program UHBD,[25] along with a term proportional to molecular surface area to account for the nonpolar solvation energy; the constant of proportionality is set to 0.006 kcal/mol/Å^2.[10,26] The interior and solvent dielectric constants set are to 1 and 80, respectively; dielectric cavity radii are set to the CHARMM22 van der Waals radii; and the dielectric boundary is defined by the Richards surface[27] with a water-sized probe sphere of radius 1.4 Å. Ionization states of the Rensing and Schrader receptor are as described in the experimental publication.[17]

3.2. *De novo* Receptor Design. Novel designs for RGD receptors were generated by an automatic construction procedure, named ConCept, that will be detailed in a separate publication. In brief, candidate receptor structures were assembled *in silico* from a set of chemical fragments comprising *N*-methyl-acetamide, *N*-methyl-formamide, acetaldehyde, acetamide, acetone, formic acid, acetic acid, ammonia, benzene, bromine, chlorine, dimethylamine, dimethyl ether, ethane, fluorine, formaldehyde, iodine, methane, methanol, methyl acetate, methyl formate, methylamine, phosphoric acid, propane, sulfonic acid, urea, and water. Chemical links were formed by deleting one hydrogen atom from each fragment and forming a single bond between the parent atoms associated with each fragment.

Receptor design is initiated by generating a stable conformation of the targeted ligand with the M2 method (Section 3.1) and then identifying interaction sites around it suitable for the placement of hydrogen bond donors and acceptors, as well as hydrophobic groups. One or more initial chemical fragments are then positioned at these sites to form an initial generation of receptors. These are then expanded and evolved through a series of generations, each comprising ∼200 structures. A new generation of receptors is formed by bringing forward the best receptors without change and by adding new fragments to the others to create new structures.

After a new fragment is added to a growing receptor structure it is rotated in 30° steps about the new single bond linking it to the rest of the receptor, and all conformations without severe steric overlaps are kept. Conformations with severe steric overlaps are candidates for formation of additional single bonds, as follows. First, when two hydrogen atoms collide but do not overlap severely, the hydrogens are deleted and a methylene is inserted to bridge the structures. Second, when two non-hydrogen atoms, each bonded to at least one hydrogen atom, bump but do not overlap severely, a hydrogen atom is deleted from each and a new single bond is added to join the atoms. Finally, when two non-hydrogen atoms overlap severely, one is deleted, its bonds are reconnected to the other parent atom, and hydrogen atoms are added or deleted to satisfy valence requirements.

Once the fragments are added and their various conformations and additional single bonds have been formed, single-atom changes are made to further optimize the receptors. This is done by categorizing each non-hydrogen atom in the receptor as a hydrophobe, a hydrogen-bond donor, or a hydrogen-bond acceptor. Then each heavy atom near a ligand interaction site (see above) is checked for congruity with the site; e.g., a donor atom should lie near a donor site. Incongruous atoms are candidates for mutation to a more suitable element.

All of the resulting receptor structures are relaxed by 50 steps of conjugate gradient energy minimization, with the target ligand held fixed. The set of receptors generated in this way is pruned back to a generation size of ∼200, based upon a rapid estimate of binding affinity for the targeted ligand. (See next paragraph.) New generations of receptors are formed until 10 generations are completed.

(18) Chang, C.-E.; Gilson, M. K. *J. Comput. Chem.* **2003**, *24*, 1987−1998.
(19) Chen, W.; Huang, J.; Gilson, M. K. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1301−1313.
(20) Chang, C.-E.; Potter, M. J.; Gilson, M. K. *J. Phys. Chem. B* **2003**, *107*, 1048−1055.
(21) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(22) Gilson, M. K.; Gilson, H. S. R.; Potter, M. J. *J. Chem. Inf. Comput. Sci.* **2003**, *43*, 1982−1997.
(23) VeraChem LLC, Germantown, MD.
(24) Qiu, D.; Shenkin, P. S.; Hollinger, F. P.; Still, W. C. *J. Phys. Chem.* **1997**, *101*, 3005−3014.
(25) Davis, M. E.; Madura, J. D.; Luty, B. A.; McCammon, J. A. *Comput. Phys. Commun.* **1991**, *62*, 187−197.
(26) Friedman, R. A.; Honig, B. *Biophys. J.* **1995**, *69*, 1528−1535.
(27) Richards, F. M. *Annu. Rev. Biophys. Bioeng.* **1977**, *6*, 151−176.

Table 1. Calculated and Experimental Standard Free Energies of Binding ΔG° for the Synthetic Receptor of Rensing and Schrader with RGD, cyclo(RGDfV), and a Nonpeptidic RGD Mimetic (kcal/mol, Standard Concentration 1 mol/L), along with Calculated Changes in Boltzmann Averaged Energy Components and Configurational Entropy[a]

	RGD tripeptide	Cyclo(RGDfV)	Benzamidine 6
ΔG°_{expt}	$-4.27 \pm 5\%$	$-3.91 \pm 14\%$	no binding detected
ΔG°_{calc}	−5.19	−3.11	−0.66
$\Delta\langle U + W\rangle$	−26.92	−25.55	−18.98
$-T\Delta S_{config}$	21.73	22.44	18.32
$\Delta\langle U_{vdw}\rangle$	−2.85	−16.16	−0.63
$\Delta\langle U_{Coul}\rangle$	−116.03	−39.97	−92.12
$\Delta\langle W_{np}\rangle$	−1.55	−2.30	−1.06
$\Delta\langle W_{elec}\rangle$	93.03	30.91	74.62
$\Delta\langle U_{val}\rangle$	0.50	1.963	0.22
$\Delta\langle U_{Coul} + W_{elec}\rangle$	−23.01	−9.06	−17.50

[a] $\Delta\langle U_{vdw}\rangle$: change in mean van der Waals energy. $\Delta\langle U_{Coul}\rangle$: change in mean Coulombic energy. $\Delta\langle W_{np}\rangle$: change in mean nonpolar solvation energy. $\Delta\langle W_{elec}\rangle$: change in mean electrostatic solvation energy. $-T\Delta S_{config}$: free energy contribution from change in configurational entropy. $\Delta\langle U_{val}\rangle$: change in sum of mean bond, angle, and torsional energies. $\Delta\langle U_{Coul} + W_{elec}\rangle$: change in mean electrostatic energy (sum of Coulombic and electrostatic solvation terms).

This evolutionary design method generates thousands of candidate receptor designs, so a rapid method of ranking them is required. The binding free energy is approximated as

$$\Delta G \approx \Delta G_{val} + \Delta G_{vdw} + \Delta G_{Coul} + \Delta G_{GB} + \Delta G_{np} \qquad (2)$$

where the terms correspond respectively to the changes in valence energy (bond stretches, angle bends, and dihedral rotations); van der Waals interactions; Coulombic interactions; electrostatic solvation, estimated with a generalized Born model;[24,28] and nonpolar solvation, estimated as proportional to surface area.[26] Energy parameters are drawn from the CHARMM force field[29] combined with VC/2004 atomic partial charges,[22] and generalized Born cavity radii and the nonpolar solvation coefficient are as previously described.[10] The change upon binding is based upon the approximation that both the ligand and the receptor remained in a fixed conformation upon dissociation.

Multiple design runs were executed, and the highest ranked receptor from each run, based upon ΔG in eq 2, was further assessed with the M2 algorithm (section 3.1). In some cases, the results of the M2 calculations were used to guide additional modifications of an initial receptor design, as described in Results.

4. Results

4.1. Analysis of Rensing and Schrader's synthetic RGD receptor. Table 1 compares experimental and calculated standard free energies of binding of Rensing and Schrader's synthetic receptor for the RGD tripeptide, a cyclic peptide containing the RGD motif (cyclo(RGDfV)), and the nonpeptidic RGD mimetic Benzamidine 6.[16] The calculations for RGD and cyclo(RGDfV) agree extremely well with experiment: the computed free energies are within 1.0 kcal/mol. In addition, the weak binding free energy of −0.66 kcal/mol calculated for Benzamidine 6 is consistent with the absence of binding observed experimentally. The calculations also provide information on the conformational preferences of the free and bound molecules and on the contributions of various energy terms to the overall free energy of binding, as now discussed.

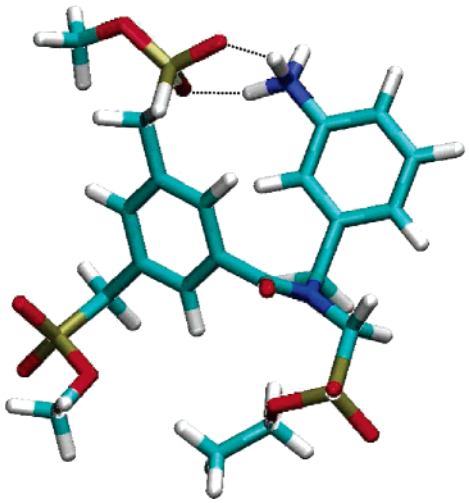

Figure 2. Most stable conformation of Rensing and Schrader's RGD receptor, based upon M2 calculations. Dashed lines: salt bridge between phosphate and ammonium groups. Cyan, carbon; red, oxygen; blue, nitrogen; brown, phosphorus; white, hydrogen. Molecular graphics here and in all other figures were generated with the program VMD.[49]

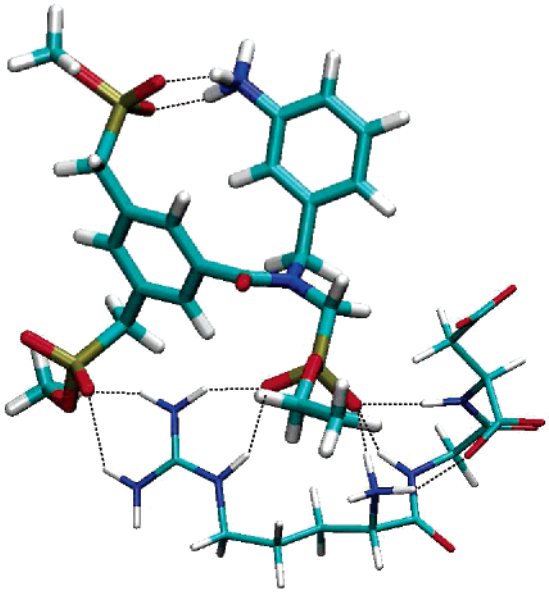

Figure 3. Most stable conformation found for complex of RGD (thin tubes) with Rensing and Schrader's receptor (thick tubes). Dashed lines highlight ionic interactions discussed in text.

4.1.1. Free Receptor. The 10 most stable conformations have free energies within 1.1 kcal/mol and are structurally similar to each other (Figure 2). In particular, they all possess a salt-bridge between one of the two chemically equivalent phosphates and the aminobenzene moiety, with oxygen−hydrogen distances within about 1.7 Å. The same salt-bridge is observed in the bound complexes, as noted in the following sections.

4.1.2. Receptor and RGD Peptide. The most stable conformation of the receptor is shown in Figure 3; again, all 10 of the most stable conformations (free energies within 0.8 kcal/mol) are similar. The interface is a zipperlike interdigitation of ions of alternating sign, comprising one of the receptor's phosphates, the peptide's guanidinium, another receptor phosphate, the peptide's terminal ammonium group, and finally the peptide's (sic) C-terminal carboxylate. The second phosphate in this sequence furthermore accepts hydrogen bonds from the

(28) Gilson, M. K.; Honig, B. *J. Comput. -Aided. Mol. Des.* **1991**, *5*, 5−20.
(29) CHARMm, version 22; Molecular Simulations Inc.: Waltham, MA., 1992.
(30) Rensing, S.; Schrader, T. *Org. Lett.* **2002**, *4*, 2161−2164. Supporting Information.

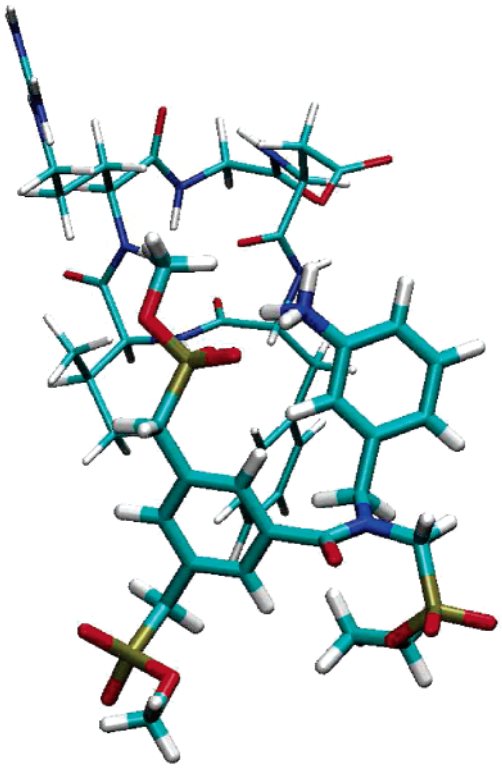

Figure 4. Most stable conformation of the complex of cylic RGD analogue cyclo(RGDfV) (thin tubes) with Rensing and Schrader's receptor (thick tubes).

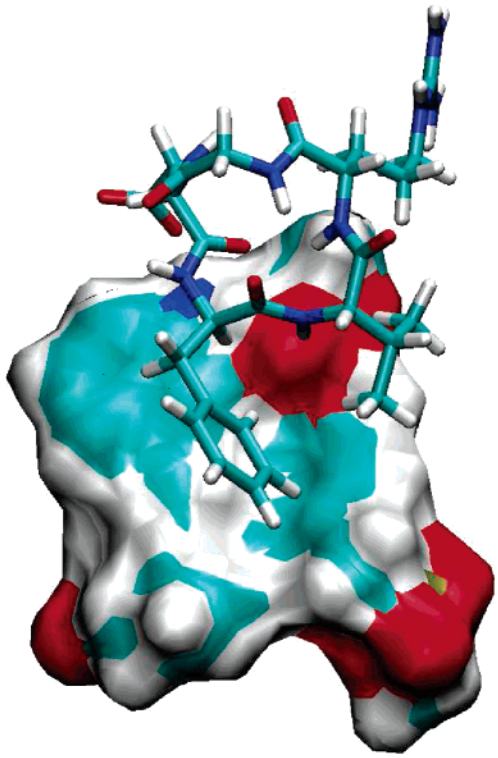

Figure 5. Alternate view of most stable conformation of the complex of cylic RGD analogue cyclo(RGDfV) (tubes) with Rensing and Schrader's receptor (solid).

peptide's two amide nitrogens. Three ionic groups do not participate in the zipper: the peptide's aspartyl side chain projects into solution, and as in the free receptor, the receptor's third phosphate forms a salt-bridge with the receptor's aminobenzene moiety. The formation of multiple stabilizing guanidinium-phosphate interactions is consistent with the designers' expectations, but the absence of close contacts between the carboxyl groups of the RGD peptide and the cationic groups of the receptor is not. However, the present model is consistent with the titratable NMR chemical shifts at the methylene and main-chain hydrogens of the aspartyl residue.[30]

As shown in Table 1, the extensive ionic interactions lead to a Coulombic stabilization of −116 kcal/mol, which is only partly balanced by an electrostatic desolvation penalty of +93 kcal/mol, leaving a net electrostatic stabilization of −23 kcal/mol. This in turn is balanced by a nearly equal loss in configurational entropy upon binding, +22 kcal/mol. Adding van der Waals interactions and nonpolar solvation changes leads to a standard binding free energy of −5.2 kcal/mol, in excellent agreement with the experimental result −4.3 kcal/mol.

4.1.3. Receptor and Cyclic RGD Derivative. Similar calculations predict that the conformation of the Rensing and Schrader receptor bound to the cyclic peptide resembles that which it adopts when bound to RGD. (Compare Figure 4 with Figure 3.) However, adding the hydrophobic Val and Phe residues to the peptide and blocking its ionic chain termini causes the ligand to bind at a different location on the receptor and to engage primarily in nonpolar, rather than polar, interactions. As highlighted in Figure 5, the receptor possesses a hydrophobic surface which receives the phenyl and valyl side chains of the ligand, while a receptor ammonium forms a salt bridge with the ligand's Asp side chain. Interestingly, the new binding mode differs from that envisioned by Rensing and Schrader (Figure 6 in ref 17): the Val and Phe side chains were thought to remain fully solvated, with the carboxylate and guanidinium groups forming salt bridges to the ammonium and phosphate groups of the receptor. In accord with the structural picture, electrostatic interactions play a much smaller role here than for the RGD peptide, and the van der Waals and nonpolar terms contribute much more. (See Table 1.) There is still a large loss in configurational entropy upon binding, and the overall energy balance leaves a somewhat weaker affinity relative to RGD. Again, the calculated free energy of binding is in excellent agreement with the experimental value of −3.9 kcal/mol.

4.1.4. Receptor and Benzamidine RGD Mimetic. The conformation of the receptor bound to the RGD mimetic Benzamidine 6 is again predicted to resemble its conformation with RGD; compare Figure 6 with Figure 3. The guanidinium moiety of the mimetic is chelated by two receptor phosphate groups, but the ligand's extended conformation makes it impossible for the receptor to form another salt bridge with the ligand's carboxylate group. Energy analysis (Table 1) shows strong electrostatic interactions but minimal nonpolar and van der Waals interactions. The loss in configurational entropy is substantial, though less than that for the other ligands. The net binding free energy is computed to be near zero, consistent with the absence of binding observed experimentally.

4.2. Novel Receptor Designs. The automated design algorithm took about a day to construct and preliminarily evaluated on the order of 20 000 candidate receptor designs. About 25 designs were selected based upon their initial energy evaluations (eq 2) and inspection for complementarity with RGD and for synthesizability. The M2 algorithm was used to compute the affinities of these designs for the RGD peptide, and four

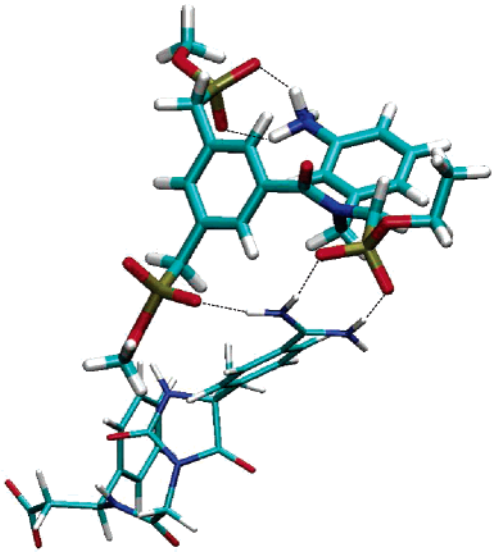

Figure 6. Most stable conformation computed for benzamidine RGD mimetic (thin tubes) with receptor of Rensing and Schrader (thick tubes). Interactions mentioned in the text are highlighted with dashed lines.

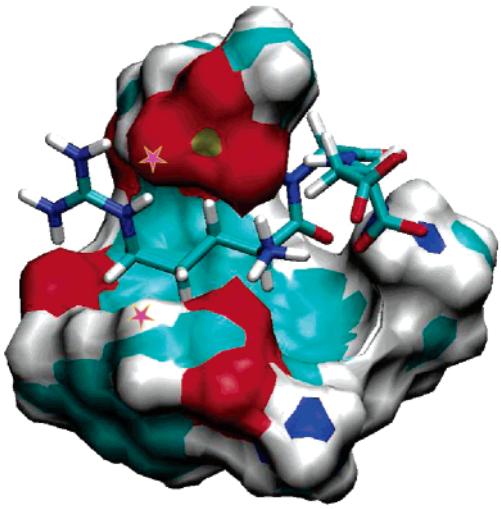

Figure 7. Most stable conformation computed for receptor L1 (surface) with bound RGD (tubes). Stars indicate sites for construction of bridge to generate receptor Bridge1.

were predicted to be of relatively high affinity: G1 and G2, which had been initialized with a single benzene fragment; and L1 and L2, initialized with three fragments, benzene, ammonia, and phosphoric acid. The M2 results then guided changes to three of these designs aimed at producing receptors with greater affinity. G1 yielded G1m1; G2 yielded G2m1; and L1 yielded L1m1, L1m2, Bridge1, and Bridge 2. Figure 1 shows the 10 designed receptors; there is no guarantee that these receptors are easy to synthesize, but they appear chemically reasonable.

4.2.1. Maximization of Affinity. Five of the new designs are predicted to bind significantly more strongly than the Rensing and Schrader receptor, even allowing for approximately 1 kcal/mol uncertainty in the calculations. The highest affinity receptor from the automated design algorithm is L1, at −6.6 kcal/mol. The most stable conformations computed for its complex with RGD (Figure 7) possess good steric and electrostatic complementarity, the receptor's four arms forming a cavity which holds the peptide. We sought to further increase affinity by using the *de novo* design software ConCept to build a link

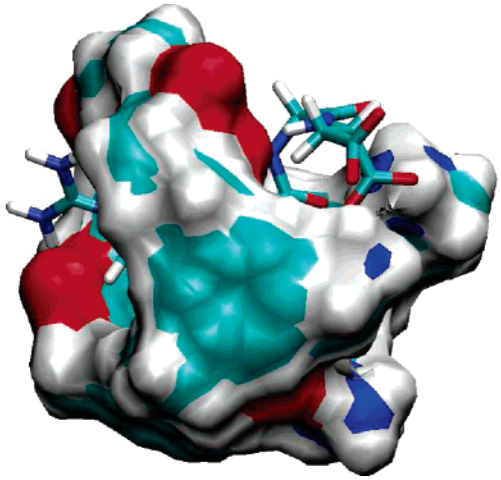

Figure 8. Most stable conformation computed for receptor Bridge1 (surface) with bound RGD (tubes).

betwen selected atoms in L1 (Figure 7). The best outcome is receptor Bridge1, a macrocycle with an added benzylamine moiety (Figure 1). Further cross-linking to form Bridge 2 led to weaker affinity.

Figure 8 shows the most stable conformation computed for Bridge1 with RGD. The conformation strongly resembles that of the L1 complex except for the added macrocyclic bridge. The bridge provides new electrostatic interactions between the added ammonium group and the terminal and aspartate carboxylates of the RGD peptide, and further stabilization from van der Waals and nonpolar solvation terms (Table 2), but loses 5 kcal/mol due to configurational entropy and also incurs additional valence strain. The result is a predicted binding free energy of −7.5 kcal/mol, the largest computed in this study.

It was surprising that Bridge1 lost more configurational entropy upon binding than Link1, because Bridge1's macrocyclic ring was expected to reduce the flexibility and hence the entropy of the free receptor. In fact, the calculations do indicate that free Link1 has 4 additional conformations within thermal energy (RT = 0.6 kcal/mol) of its most stable conformation, whereas Bridge1 has none, consistent with expection. Moreover, when affinities are recomputed based upon only the single most stable conformation of the free and bound receptors, the affinities for L1 and Bridge1 are found to be within 0.1 kcal/mol. Only when all conformations are included does the difference rise to 1.1 kcal/mol, as per the data in Table 2. Thus, the macrocycle of receptor Bridge1 does promote binding by reducing the loss of accessible conformations upon binding, as anticipated. However, Bridge1 still loses more configurational entropy upon binding than does L1.

The explanation is that the individual bound conformations of Bridge1 have less entropy, relative to the free receptor, than those of L1: each energy well becomes narrower for Bridge 1 than for L1. This can be seen by comparing the entropies of the most stable bound and free conformations of the two receptors: the rise in $-T\Delta S°$ computed in this way is found to be 7 kcal/mol greater for Bridge1 than for Link1. Thus, although Bridge1 does not lose as many distinct low-energy conformations upon binding as L1, its bound conformations are held more snugly and thus generate a greater entropy penalty.

4.2.2. Per Atom Efficiency. The binding free energy per atom of a receptor provides another measure of receptor quality. The

Table 2. Calculated Standard Free Energies of Binding ΔG°_{calc} for RGD with Rensing and Schrader's Receptor (RS) and Designed Receptors Bridge 1, L1m2, L1, L1m1, L2, Bridge 2, G1m1, G2, G2m1, and G1, in Order of Decreasing Computed Affinity, along with Calculated Changes in Boltzmann Averaged Energy Components, Configurational Entropy and Surface Area (A^2), and Measures of Efficiency; See Table 1 for Symbols

	RS	Bridge1	L1m2	L1	L1m1	L2	Bridge2	G1m1	G2	G2m1	G1
ΔG°_{calc}	−5.19	−7.54	−6.82	−6.64	−6.58	−6.31	−5.93	−5.0	−4.89	−4.74	−4.71
$\Delta\langle U + W\rangle$	−26.92	−40.94	−25.68	−34.83	−25.58	−31.26	−35.49	−26.0	−28.89	−31.87	−27.2
$-T\Delta S^\circ_{config}$	21.73	33.4	18.89	28.19	19.00	24.95	29.56	21.0	24.00	27.14	22.50
$\Delta\langle U_{vdw}\rangle$	−2.85	−16.19	−13.28	−13.69	−6.91	−10.64	−12.63	−7.07	−12.62	−18.49	−7.72
$\Delta\langle U_{Coul}\rangle$	−116.03	−196.43	−62.24	−78.06	−71.80	−120.75	−143.14	−105.0	−72.19	−110.14	−115.5
$\Delta\langle W_{np}\rangle$	−1.55	−3.71	−1.93	−2.29	−1.63	−2.25	−2.73	−1.84	−2.06	−3.02	−1.84
$\Delta\langle W_{elec}\rangle$	93.03	172.93	50.66	60.36	54.99	100.72	119.89	83.3	58.80	97.17	96.15
$\Delta\langle U_{val}\rangle$	0.50	2.46	1.12	−1.15	−0.23	1.66	3.12	2.29	−0.81	2.60	1.69
$\Delta\langle U_{Coul} + W_{elec}\rangle$	−23.01	−23.50	−11.59	−17.70	−16.81	−20.03	−23.26	−15.15	−13.39	−12.97	−19.34
$N_{heavyatoms}$	36	60	53	48	35	40	72	24	44	44	32
per atom efficiency	0.14	0.13	0.13	0.14	0.19	0.16	0.08	0.21	0.11	0.11	0.15
energy efficiency	0.19	0.18	0.27	0.19	0.26	0.20	0.17	0.19	0.19	0.15	0.17
electrostatic efficiency	0.20	0.12	0.19	0.23	0.23	0.17	0.16	0.14	0.19	0.12	0.17
change in SASA	−258	−618	−322	−382	−272	−375	−455	−307	−343	−503	−307
interfacial efficiency	0.020	0.012	0.021	0.017	0.024	0.017	0.013	0.016	0.014	0.009	0.015

per atom efficiency of the receptors studied here varies nearly 3-fold, from 0.08 kcal/mol/atom for Bridge2 to 0.21 kcal/mol/atom for G1m1; the latter was derived from the automatically designed receptor G1 by substituting a methyl group for a benzoic acid group that was computed to project into solution instead of interacting with RGD. The per atom efficiency of the Rensing and Schrader receptor is about average, at 0.14 kcal/mol/atom.

4.2.3. Configurational Entropy and Energy Efficiency. The computed change in configurational entropy upon binding imposes remarkably large free energy penalties of 19−30 kcal/mol, especially in comparison to the total binding free energies of −4.7 to −7.5 kcal/mol (Table 2). The changes in configurational entropy correlate with the changes in mean energy ($\langle U + W\rangle$), as shown in the upper graph of Figure 9, and the data fall along the same line seen previously for aqueous cyclodextrins and synthetic receptors in chloroform. Fitting the combined data yields $T\Delta S^\circ \approx -0.83\Delta\langle U + W\rangle - 0.51$, which, with eq 1, yields $\Delta G^\circ \approx 0.17\ \Delta\langle U + W\rangle - 0.51$. The lower graph of Figure 9 shows that this approximation is not very accurate. This implies that the deviations from the fit in the upper portion of the figure are substantial on the scale of the binding energies plotted in the lower portion of the figure and, thus, that some receptors pay a proportionally smaller entropy penalty than others. The degree to which a receptor's binding energy survives entropic compensation can be expressed as its energy efficiency, the ratio of its binding free energy to the change in average energy, $\Delta\langle U + W\rangle$. This quantity ranges from a low of 0.15 for G2m1 to 0.27 for L1m2, with a mean of 0.19. (See Table 2.) The energy efficiency of the Rensing and Schrader receptor is average, as is that of design Bridge 1.

4.2.4. Solvation and Electrostatic Efficiency. Favorable Coulombic interactions formed on binding are largely compensated by the cost of desolvating polar groups: indeed, a powerful negative correlation is observed between the change in mean Coulombic energy upon binding and the change in mean electrostatic solvation energy, for the highly polar RGD systems studied here and for much less polar cyclodextrin-ligand systems studied previously.[10] (See Figure 10 (top).) To our knowledge, the strength and linearity of this correlation has not been noted previously. Linear regression yields $\Delta\langle W_{elec}\rangle \approx -0.84\Delta\langle U_{Coul}\rangle$. Hence the net change in electrostatic energy can be approximated by $\Delta\langle U_{Coul} + W_{elec}\rangle \approx 0.16\Delta\langle U_{Coul}\rangle$. Figure 10 (bottom) shows that this approximation works fairly well, but the scatter is large on the scale of the binding free energies to which these terms contribute. It is not surprising that the errors are smaller for the cyclodextrins, since electrostatic forces are weaker for these less polar systems.

The receptor−ligand systems which fall most below the line in the bottom graph of Figure 10 are the ones which gain the greatest net electrostatic stabilization for a given Coulombic contribution. The variation is substantial: for example, the cluster of points with $\Delta\langle U_{Coul}\rangle \approx -110$ kcal/mol shows net electrostatic energies that vary over about 10 kcal/mol. This variation directly affects the bottom line free energy of binding,

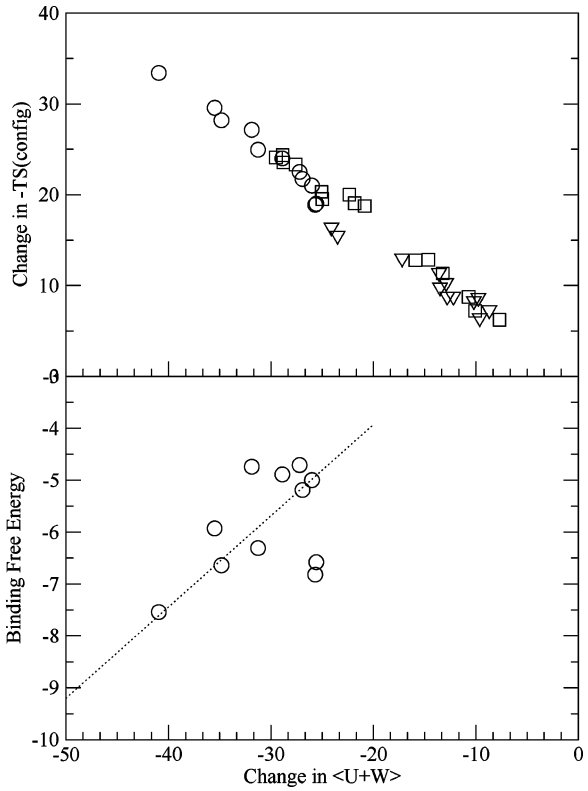

Figure 9. Change in contribution of configurational entropy (top) and binding free energy (bottom) versus change in mean energy ($\langle U + W\rangle$) for all RGD receptors listed in Table 2 (○), for aqueous cyclodextrin−guest complexes[10] (□), and for other host−guest systems in chloroform[9] (▽). All quantities are computed and are in units of kcal/mol.

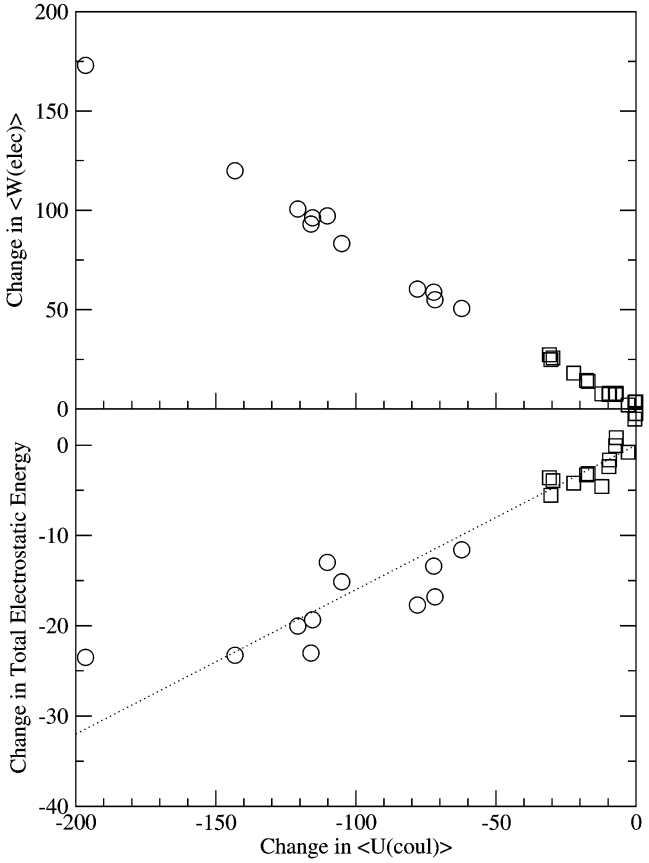

Figure 10. Correlation analysis of computed changes in electrostatic energy components upon binding for RGD receptors studied here (○) and cyclodextrin−guest systems studied previously[10] (□). All systems are aqueous, and energies are Boltzmann averages, in kcal/mol. Top: Change in electrostatic solvation energy versus change in Coulombic energy. Bottom: Change in sum of solvation and Coulombic energies versus change in Coulombic energy.

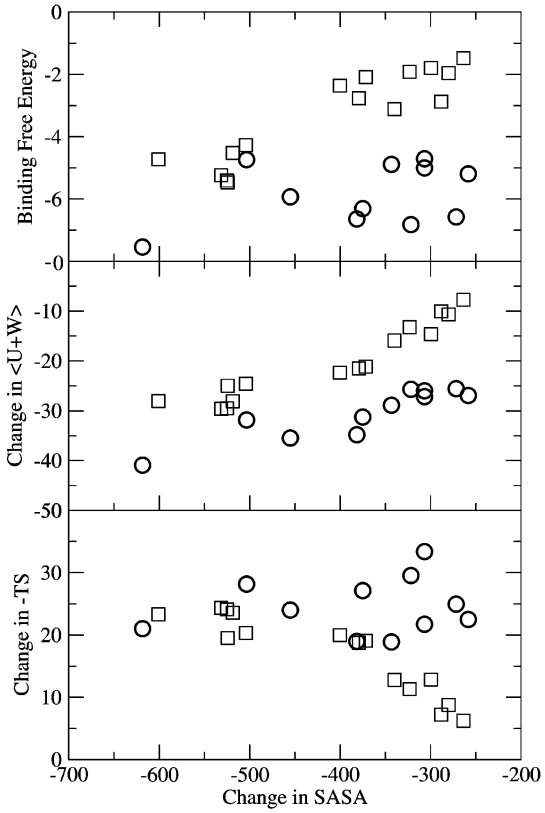

Figure 11. Relationships between change in solvent-accessible surface area (Angstroms) and computed binding free energy (top), change in Boltzmann-averaged total energy ($\langle U + W \rangle$) (middle) and change in configurational entropy ($-T\Delta S^{\circ}_{\text{config}}$) (bottom) upon binding, for RGD receptors considered here (○) and for cyclodextrin systems reported previously.[10] Energies in kcal/mol, surface areas in Å^2.

so it is of interest to define a corresponding measure of efficiency, the total electrostatic driving force for binding divided by the change in the purely Coulombic energy: $\Delta\langle U_{\text{Coul}} + W_{\text{elec}}\rangle/\Delta\langle U_{\text{Coul}}\rangle$. This quantity provides information on how much of the designed-in Coulombic interactions are expressed in the total electrostatic energy after accounting for the electrostatic solvation penalty. As shown in Table 2, this quantity varies from 0.12 for G2m1 and Bridge 1 up to 0.23 for L1 and L1m1, with a mean of 0.17. The Rensing and Schrader design is slightly better than average, by this measure.

4.2.5. Surface Area and Interfacial Efficiency. The loss in solvent-exposed surface area upon binding correlates roughly with affinity across a range of systems,[4] but little or no such correlation is evident within the comparatively narrow series of RGD systems studied here (Figure 11, top). Although a correlation is evident for the previously studied cyclodextrin systems,[10] there is no convincing correlation for the combined data set (Figure 11, top). This result is consistent with data summarized in the prior study (Tables 3, 4, and 6 of ref 4), which show little or no correlation between buried surface area and affinity for the association of α- and β-cyclodextrin with various ligands or for a set of catalytic antibodies.

Interestingly, the change in energy upon binding, $\Delta\langle U + W\rangle$, correlates fairly strongly with surface area within each receptor set, as shown in the middle graph of Figure 11. The change in configurational entropy shows no correlation at all with buried surface area for the RGD systems, but there is a rather strong negative correlation for the cyclodextrin systems (Figure 11, bottom). However, the combined data sets, again, show little or no correlation between buried surface area and either $\Delta\langle U + W\rangle$ or $-T\Delta S^{\circ}_{\text{config}}$.

Given the expectation that buried surface area should correlate with affinity, one may characterize a receptor according to how effectively it uses its buried surface, defining its interfacial efficiency as the ratio of binding free energy to the change in surface area upon binding. Values for the RGD system average 0.016 kcal/mol/Å^2 and range from 0.009 for G2m1 up to 0.024 for L1m1 (Table 2). This designed receptor also scores well on both energy and electrostatic efficiency; in effect, it makes good use of each square angstrom of buried surface area by paying proportionally low entropic and solvation penalties. More generally, the interface efficiency correlates with both the entropic and electrostatic efficiencies, as shown in Figure 12.

5. Discussion

5.1. Modeling Rensing and Schrader's Receptor. This study was inspired in large part by the pioneering work of Rensing and Schrader, who created a compact synthetic receptor which binds the RGD peptide with a standard free energy change of −4.3 kcal/mol. The M2 method yields strikingly accurate computed affinities for this receptor, agreeing with experiment to within about 1 kcal/mol for binding of RGD, as well as a cyclic variant of RGD and a nonpeptidic RGD-mimetic. These results are significant because the M2 method had not hitherto

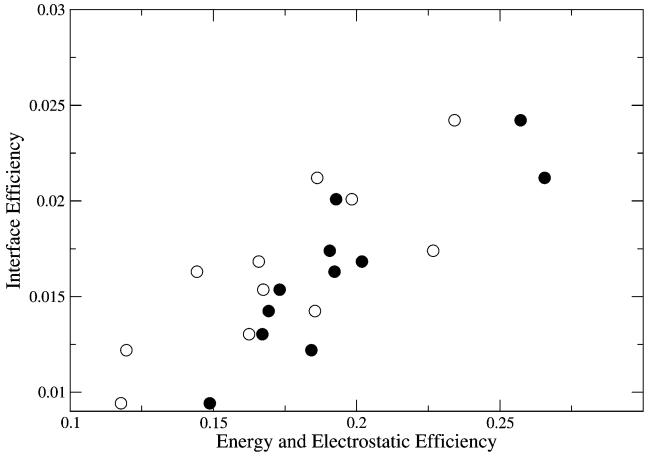

Figure 12. Correlation of interfacial efficiency with energy efficiency (●) and electrostatic efficiency (○).

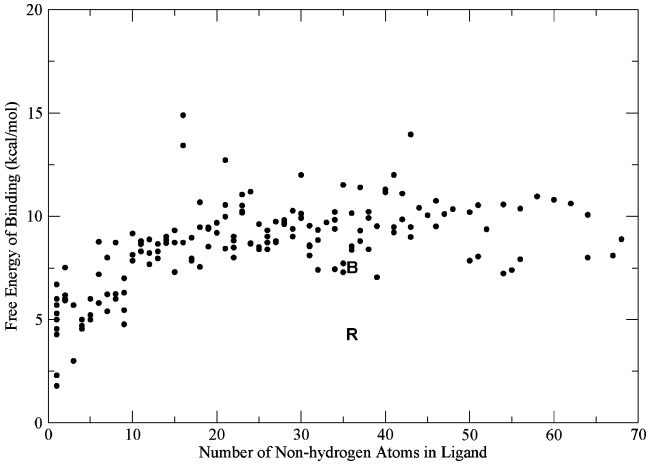

Figure 13. Standard free energy of binding (times −1) versus number of non-hydrogen atoms of ligand, for proteins (●) and for RGD receptors. "R": Rensing and Schrader's receptor. "B": Bridge1, the best design from this study.

been tested for such highly ionic compounds; previous studies yielded similar accuracies for cyclodextrins in water[10] and synthetic receptors in chloroform.[9]

The most stable conformations of the bound complexes predicted here differ from those anticipated by Rensing and Schrader but are consistent with the available NMR data. Interestingly, completely different binding modes are predicted for RGD and its cyclic variant, the former relying primarily upon ionic interactions, and the latter upon hydrophobic interactions. This result seems reasonable, given that the cyclic peptide has two fewer ionic groups (the termini) and two additional nonpolar groups (the Phe and Val side chains) and that the edge of the receptor is highly ionic while its faces are nonpolar. These structural insights would be valuable if one wished to use the Rensing and Schrader receptor as a starting point for the design of even more highly optimized receptors.

The receptor's predicted chelation of the guanidinium group of RGD with two phosphate moieties resembles the chelation of the guanidinium of the cyclic RGD derivative by two aspartyl side chains in Integrin $\alpha V\beta 3$.[31] Note, however, that divalent cations play a key role in the interaction of integrins with the carboxylate group of the RGD motif, but the Rensing and Schrader receptor was studied in a solution lacking divalent cations. Interestingly, whereas hydrophobic interactions are predicted to be important in the association of the Rensing and Schrader host with the cyclic peptide, hydrophobic interactions play little role in the association of a similar cyclic peptide with Integrin $\alpha V\beta 3$.[31] Presumably, the interaction is dominated by the highly organized and correctly spaced subsites for the guanidinium and carboxylate side chains.

The affinity of the Rensing and Schrader receptor for RGD falls significantly below the range of affinities expected for protein−ligand binding,[32] when the ligand is the size of RGD, as shown in Figure 13. In addition, the M2 calculations suggest that much of the receptor does not interact directly with RGD (section 4.1.2). These observations motivated a search for receptors that would bind more strongly or more efficiently.

5.2. New Receptor Designs. Several of the designed receptors are predicted to bind RGD significantly more tightly than the Rensing and Schrader receptor. Encouragingly, the affinity of Bridge1 for RGD, −7.5 kcal/mol, rivals the affinities of proteins for ligands of similar size (Figure 13).[32] Other new designs also are predicted to bind RGD more tightly than that of Rensing and Schrader or to bind as tightly but more efficiently on a per atom basis. These results are promising, but proteins often achieve binding free energies of about −12 kcal/mol for ligands the size of RGD (Figure 13) and such affinities lie beyond the historical limit for aqueous host−guest systems listed in a recent compilation.[4] Searching for a synthetic receptor which generates affinities this high, or for the reason one cannot exist, is of fundamental interest, and the following subsections address the challenge of achieving high affinities with synthetic receptors.

5.3. Concepts in Receptor Design. A central observation of the present study is that energies that drive binding are strongly balanced by proportionate repulsive ones. These compensations pose a major obstacle to the design of high affinity synthetic receptors. Thus, for the RGD systems, strengthening attractive Coulombic interactions by 10 kcal/mol generates an electrostatic desolvation penalty of about 8 kcal/mol (Figure 10). (The importance of electrostatic desolvation has been emphasized previously; see, e.g., refs 33−36.) Similarly, strengthening the full energy, $(\Delta\langle U + W\rangle)$, by 10 kcal/mol engenders an entropy penalty of about 8 kcal/mol (Figure 9). This entropic compensation is similar to the experimental phenomenon of entropy−enthalpy compensation (e.g., refs 4, 9, 10, 37−40) and presumably reflects the fact that stronger attractive forces produce a more rigid complex.

One way to enhance affinity is to design receptors that form more extensive contacts with the ligand, as evidenced by the broad correlation between surface area and affinity.[4] Even

(31) Xiong, J.-P.; Stehle, T.; Zhang, R.; Joachimiak, A.; Frech, M.; Goodman, S. L.; Arnaout, M. A. *Science* **2002**, *296*, 151−155.
(32) Kuntz, I. D.; Chen, K.; Sharp, K. A.; Kollman, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9997−10002.
(33) Honig, B. H.; Hubbell, W. L. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 5412−5416.
(34) Sheinerman, F. B.; Honig, B. *J. Mol. Biol.* **2002**, *318*, 161−177.
(35) Gilson, M. K.; Honig, B. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 1524−1528.
(36) Hendsch, Z. S.; Tidor, B. *Protein Sci.* **1994**, *3*, 211−226.
(37) Lumry, R.; Rajender, S. *Biopolymers* **1970**, *9*, 1125−1227.
(38) Krug, R.; Hunter, W.; Grieger, R. *Nature* **1976**, *261*, 566−567.
(39) Guo, Q.; Zheng, X.; Ruan, X.; Luo, S. H.; YC, Y. L. *J. Inclus. Phenom. Mol.* **1996**, *26*, 233−241.
(40) Sharp, K. *Protein Sci.* **2001**, *10*, 661−667.

though entropic losses cancel much of the resulting energetic benefit, a net gain can still be achieved. However, this approach will lead to relatively large and complex receptors. Moreover, increasing the surface area buried upon binding is by no means guaranteed to increase affinity (section 4.2.5), since the change in mean energy upon binding, $\Delta\langle U + W\rangle$, correlates only weakly with surface area. Moreover, there is no compelling physical reason that the loss in configurational entropy should correlate with buried surface area. In general, the amount of buried surface area may impose an upper limit on affinity, but it does not impose a lower limit, so a limited focus on maximization of interfacial surface area does not represent a complete strategy for maximizing affinity.

A more elegant concept would be to design receptors that fall off the compensation curves shown here, i.e., receptors that pay proportionally smaller entropy and electrostatic desolvation penalties. Recently developed methods of optimizing the balance between electrostatic attractions and desolvation (see, e.g., refs 41−44), as well as the method of sensitivity analysis,[45−48] may be useful for the electrostatic part of the problem, especially for highly ionic ligands such as the RGD peptide. Entropic losses will presumably be minimized by maximizing preorganization without sacrificing the geometric precision of receptor−ligand complementarity. However, it is not clear that we adequately understand the rules required to guide such designs.

In fact, one traditional approach to reducing the entropy loss upon binding, cyclization, is strikingly unsuccessful in the present study. Cyclizing receptor L1 to form receptors Bridge 1 and Bridge 2 was expected to reduce the entropy of the free receptor preferentially, and thus reduce the entropic penalty upon binding. However, the entropy losses rose instead (Section 4.2.1): cyclization did preferentially reduce the number of stable conformations (energy wells) in the free state, but it also disproportionately narrowed the energy wells associated with the bound conformations and the entropy penalty from this narrowing outweighed the entropy benefit from changing the numbers of energy wells. The narrowing of the energy wells in the bound state appears traceable to the snug and hence restrictive binding site of Bridge 1. The upshot is that both L1 and Bridge 1 fall squarely on the best-fit linear energy−entropy relationship shown in Figure 9. To our knowledge, it has not previously been shown that cyclization can worsen entropy losses by disproportionately restricting the bound complex, but this result does agree with a prior observation that snugness of fit correlates with entropy loss.[10] Until rules are available for predicting the entropic consequences of cyclization, computational methods such as M2 should be useful for assessing design proposals.

5.4. Novel Measures of Receptor Efficiency. A given change in receptor−ligand surface area upon binding arguably places an upper limit upon the affinity achievable by noncovalent forces. It is thus of interest to characterize receptor−ligand complexes by their binding free energy per square angstrom of surface area, the interfacial efficiency introduced in section 4.2.5, and to inquire into the determinants of this measure of receptor efficiency.

One determinant is the energy efficiency, the effectiveness of the receptor's use of the overall energy change upon binding (section 4.2.3). This quantity measures the degree to which the energetic driving forces for binding outweigh the losses in configurational entropy. Another determinant is the electrostatic efficiency (section 4.2.4), which measures the degree to which attractive Coulombic interactions outweigh the inevitable penalty in electrostatic desolvation. These new measures of receptor efficiency are reminders that maximizing receptor affinity requires more than maximizing buried surface area, overall attractive interactions, or Coulombic complementarity; it is also necessary to minimize losses in configurational entropy and electrostatic desolvation if it is to achieve proteinlike affinities. These measures of efficiency also provide a basis for considering the upper limits of affinity that are achievable within given parameters. For example, a receptor with the large surface area change of Bridge 1 and the high interfacial efficiency of L1m1 would bind RGD with a large, favorable free energy change of −14 kcal/mol. Alternatively, a receptor with the same value of $\Delta\langle U + W\rangle$ but the energy efficiency of L1m1 would bind RGD with a free energy change of −11 kcal/mol. Finally, changing the electrostatic efficiency of Bridge 1 to the value found for L1m2, while leaving the energy efficiency unchanged, leads to a binding affinity of −10 kcal/mol. There is no obvious reason receptors with such high affinity could not be devised. The present analysis thus supports the feasibility of designing targeted synthetic receptors with proteinlike affinities.

6. Acknowledgments

The authors thank Drs. Ken Houk and Hillary S. R. Gilson for thoughtful comments on the manuscript. This publication was made possible by Grant Number GM61300 from the National Institute of General Medical Sciences (NIGMS) of the NIH. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIGMS.

Supporting Information Available: Two ZIP files are provided. One contains an MDL Molfile for each of the designed receptors and for the prior design by Rensing and Schrader (RS.mol). The second contains a PDB format file with the 3D coordinates of the most stable conformation found for each receptor free in solution and then bound to RGD. Structures are included also for cyclo-RGDfV and the non-peptide RGD-mimetic studied by Rensing and Schrader. This material is available free of charge via the Internet at http://pubs.acs.org.

JA056600L

(41) Lee, L. P.; Tidor, B. *J. Chem. Phys.* **1997**, *106*, 8681−8690.
(42) Kangas, E.; Tidor, B. *J. Chem. Phy.* **1998**, *109*, 7522−7545.
(43) Sulea, T.; Purisima, E. O. *J. Phys. Chem. B* **2001**, *105*, 889−899.
(44) Kangas, E.; Tidor, B. *J. Phys. Chem. B* **2001**, *105*, 880−888.
(45) Zhang, H.; Wong, C. F.; Thacher, T.; Rabitz, H. *Protein: Struct., Funct., Genet.* **1995**, *23*, 218−232.
(46) Wong, C. F.; Thacher, T.; Rabitz, H. *Reviews in Computational Chemistry*; Wiley-VCH: New York, 1998; pp 281−326.
(47) Sims, P. A.; Wong, C. F.; McCammon, J. A. *J. Med. Chem.* **2003**, *46*, 3314−3325.
(48) Sims, P. A.; Wong, C. F.; McCammon, J. A. *J. Comput. Chem.* **2004**, *25*, 1416−1429.
(49) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graph.* **1996**, *14*, 33−38.