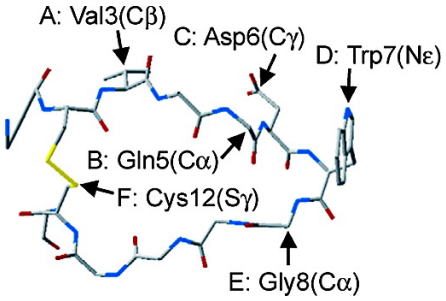
Article

Development of a Quasi-Dynamic Pharmacophore Model for Anti-Complement Peptide Analogues

Buddhadeb Mallik, and Dimitrios Morikis

J. Am. Chem. Soc., **2005**, 127 (31), 10967-10976• DOI: 10.1021/ja051004c • Publication Date (Web): 15 July 2005



Pharmacophore 1: A-C-D-F
Pharmacophore 2: A-E-F
Pharmacophore 3: A-B-E-F

More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 1 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES
Published on Web 07/15/2005

Development of a Quasi-Dynamic Pharmacophore Model for Anti-Complement Peptide Analogues

Buddhadeb Mallik and Dimitrios Morikis*

Contribution from the Department of Chemical and Environmental Engineering, University of California at Riverside, Riverside, California 92521

Received February 16, 2005; E-mail: dmorikis@engr.ucr.edu

Abstract: Three quasi-dynamic pharmacophore models have been constructed for the complement inhibitor peptide compstatin, using first principles. Uniform sampling along 5-ns molecular dynamics trajectories provided dynamic conformers that are thought to represent the entire conformational space for nine training set molecules, compstatin, four active analogues, and four inactive analogues. The pharmacophore models were built using mixed physicochemical and structural properties of residues indispensable for structural stability and activity. Owing to the size and flexibility of compstatin, one-dimensional probability distributions of intrapharmacophore point distances, angles, and dihedral angles of different analogues spread over wide and overlapping ranges. More robust two-dimensional distance−angle probability distributions for two pharmacophore models discriminated individual analogues in terms of specific distance−angle pairs, but overall failed to identify the active and the inactive analogues as two distinct groups. Two-dimensional distance−dihedral angle probability distributions in a third pharmacophore model allowed discrimination of the groups of active and inactive analogues more effectively, with the highest-activity analogue having distinct behavior. The present study indicates that more stringent structural constraints should be used for a set of structurally similar but flexible peptides, as opposed to organic molecules, to convert dynamic conformers into pharmacophore models. Flexibility is a general aspect of the structure and function of peptides and should be taken into account in ligand-based pharmacophore design. However, the discrimination of activity using multidimensional probability surfaces depends on the peptide system, the selection of the training set, the molecular dynamics protocol, and the selection of the type and number of pharmacophore points.

Introduction

Activation of the complement system provides major defense against invading foreign pathogens and bioincompatible foreign material.[1] The intrinsic characteristic of this defense mechanism is the ability of complement proteins to discriminate self from nonself. Under normal physiological conditions, various routes of this defense mechanism are precisely regulated internally by a number of structurally and functionally related proteins, called regulators of complement activation (RCA). However, in many pathological situations unregulated complement activation is possible, which can lead to damage of host tissue with implications in health.[2] This is the case in Alzheimer's disease, asthma, adult respiratory distress syndrome, burn injuries, trauma, hemolytic anemia, ischemia/reperfusion injuries, cardiopulmonary bypass surgery, rheumatoid arthritis, stroke, allotransplantation, etc. A drug that controls unregulated complement activation is a priority in drug discovery and complement research. A number of complement inhibitors have been identified; some are under different stages of clinical trials, but none are yet clinically available.[3]

A good candidate to become a drug for clinical use against unregulated complement activation is the 13-residue cyclic peptide compstatin.[4−6] Compstatin binds complement component C3 of human and primate species and inhibits complement action. Compstatin was identified in a phage-displayed random peptide library screening against C3 and was further optimized using rational design.[6−11] The sequence of compstatin with

(1) (a) Walport, M. J. *N. Engl. J. Med.* **2001**, *344*, 1058−1066. (b) Walport, M. J. *N. Engl. J. Med.* **2001**, *344*, 1140−1144. (c) Carroll, M. C. *Annu. Rev. Immunol.* **1998**, *16*, 545−568. (d) Morikis, D.; Lambris, J. D., Eds. *Structural Biology of the Complement System*; CRC Press: Boca Raton, Florida, 2005.

(2) Lambris, J. D., Holers, V. M., Eds. *Therapeutic Interventions in the Complement System*; Humana Press: Totowa, New Jersy, 2000.

(3) (a) Morgan, B. P.; Harris, C. L. *Mol. Immunol.* **2003**, *40*, 159−170. (b) Makrides, S. C. *Pharmacol. Rev.* **1998**, *50*, 59−87. (c) Sahu, A.; Lambris, J. D. *Immunopharmacology* **2000**, *49*, 133−148.

(4) (a) Morikis, D.; Lambris, J. D. In *Structural Biology of the Complement Susyem*: CRC Press: Boca Raton, Florida, 2005; pp 317−340. (b) Morikis, D.; Soulika, A. M.; Mallik, B.; Klepeis, J. L.; Floudas, C. A.; Lambris, J. D. *Biochem. Soc. Trans.* **2004**, *32*, 28−32. (c) Morikis, D.; Lambris, J. D. *Biochem. Soc. Trans.* **2002**, *30*, 1026−1036.

(5) Sahu, A.; Morikis, D.; Lambris, J. D. In *Therapeutic Interventions in the Complement System*; Lambris, J. D., Holers, V. M., Eds.; Humana Press: Totowa, New Jersy, 2000; pp 75−112.

(6) Holland, M. C. H.; Morikis, D.; Lambris, J. D. *Curr. Opin. Invest. Drugs* **2004**, *5*, 1164−1173.

(7) Sahu, A.; Kay, B. K.; Lambris, J. D. *J. Immunol*. **1996**, *157*, 884−891.

(8) Morikis, D.; Assa-Munt, N.; Sahu, A.; Lambris, J. D. *Protein Sci.* **1998**, *7*, 619−627.

(9) Sahu, A.; Soulika, A. M.; Morikis, D.; Spruce, L.; Moore, W. T.; Lambris, J. D. *J. Immunol.* **2000**, *165*, 2491−2499.

(10) Morikis, D.; Roy, M.; Sahu, A.; Troganis, A.; Jennings, P. A.; Tsokos, G. C.; Lambris, J. A. *J. Biol. Chem.* **2002**, *277*, 14942−14953.

(11) Mallik, B.; Katragadda, M.; Spruce, L. A.; Carafides, C.; Tsokos, C. G.; Morikis, D.; Lambris, J. D. *J. Med. Chem.* **2005**, *48*, 274−286.

blocking groups at the termini is Ac-Ile1-[Cys2-Val3-Val4-Gln5-Asp6-Trp7-Gly8-His9-His10-Arg11-Cys12]-Thr13-NH_2, where brackets denote cyclization through Cys2−Cys12 disulfide bond formation. The three-dimensional structure of compstatin was determined by using nuclear magnetic resonance (NMR) spectroscopy.[8,12] The structure of a major conformer of compstatin has a β-turn at the segment Gln5-Asp6-Trp7-Gly8 and a hydrophobic patch at the termini that includes the disulfide bond.[8,12] The β-turn is part of a polar face opposite the hydrophobic patch. Use of the structural information, rational optimization,[9−11] experimental combinatorial optimization,[13] and computational combinatorial optimization[14] methods identified a large number of active peptidic or peptidomimetic analogues. Structure−activity correlations using NMR spectroscopy and IC_{50} values, binding studies using surface plasmon resonance and C3 targets from various species, and thermodynamic analysis using isothermal titration calorimetry (ITC) for various compstatin analogues has allowed the proposition of hypotheses on the mode of compstatin−C3 association.[9,10,15]

A molecular dynamics simulation study of compstatin has identified the presence of five interconverting conformers with populations ranging from 9 to 44%.[16] These conformers, averaged out in the NMR study, indicate the dynamic nature of compstatin in solution and were not incorporated explicitly in previous rational and other optimization protocols. In the present study we use structural templates from molecular dynamics simulations of compstatin analogues to build a quasi-dynamic pharmacophore model, which will form the basis of a new search for alternative inhibitors. The term "quasi" denotes the use of snapshots from the molecular dynamics simulation.

A pharmacophore model represents a three-dimensional spatial arrangement for functional groups or atoms of a ligand with drug-like properties. Charged atoms, hydrophobic groups, hydrogen bond donors/acceptors, aromatic ring systems are common functional groups in ligands. The relative positions of these groups or atoms are expected to be responsible for the functionality of the ligands. Correlation of these functional groups or atoms among a series of structurally similar molecules with a range of activity is the basis of development of a ligand-based pharmacophore model. The model could be static or dynamic, depending on the nature of the data incorporated into it. Standard methods for both ligand-based[17] and receptor-based[18] development of pharmacophore models are typically limited to the use of static structures; however, the use of dynamic conformations in computational drug design has been increased recently.[19] The static model is based on the minimum-energy structures and lacks information on the dynamic behavior of molecules, which is believed to play a role in binding. Several limitations of the static models have been pointed out by Bernard et al.[20] The first limitation is the loss of the inherent dynamic motion of low-molecular mass molecules and the nature of interaction with biomolecules. Due to the thermal energy, small molecules in solution at room temperature are in random motion that promotes low energy-barrier conformational interconversion. Successful collision between the suitable conformation of a ligand with the biomolecule is the first step in drug binding. If motion is ignored, as is the case in a static model, the conformation that leads to binding might be wrongly represented in theoretical studies. The second limitation is the minimization algorithms, because the protocols and the force fields employed by different programs sample the energy landscape differently. This may result in different minimized structures, all of which sample local minima. The third limitation is that the lowest-energy structure is not necessarily the one that binds. To avoid these limitations, Bernard et al.[20] have introduced dynamic conformations from the MD trajectory of a number of nonpeptidic δ-opioid agonists and antagonists in their pharmacophore model. Statistical sampling of conformations and probability distributions of geometric parameters were used to plot distance−angle probability contours to differentiate preferred positions in two-dimensional space occupied by agonists and antagonists.

In the present study we have used the idea of Bernard et al.[20] involving the parent peptide compstatin and a set of four active and four inactive analogues of compstatin to develop a three-dimensional pharmacophore model. Compstatin and its analogues are larger and internally more flexible peptidic molecules, which contain several important locations to be selected as pharmacophore points, compared to the organic molecules of Bernard et al.[20] Except for the parent peptide compstatin, we do not have experimentally determined structures for the analogues we used in the present study. However, several compstatin analogues, including five of those we used here, were previously thoroughly characterized by NMR and were found to have structures similar to that of the parent peptide. Therefore, to perform our MD simulations, we have incorporated theoretical mutations in the structure of the parent peptide compstatin.

Methods

In this study we have considered nine analogues in total, compstatin and four active and four inactive analogues. These analogues together with their relative activities are summarized in Table 1. The five active analogues we have selected here (out of a large number of identified active analogues)[10,11,14] represent amino acid replacements that were major breakthroughs in the design optimization of compstatin. These breakthroughs were (i) acetylation of the amino terminus,[4,9,10,13] (ii) the presence of a helix promoter at position 9,[8,10,11] and (iii) the presence of an aromatic amino acid at position 4.[11,14] The five active analogues span the entire range of activity starting from the parent peptide compstatin to the highest active analogue (45-fold) with natural amino acids. The selected inactive analogues perturb major structural elements of compstatin, namely the disulfide bond and the β-turn.

We have chosen three sets of pharmacophore points shown in Figure 1. Two of these sets involve the disulfide bonded Cys2−Cys12, the

(12) Klepeis, J. L.; Floudas, C. A.; Morikis, D.; Lambris, J. D. *J. Comput. Chem.* **1999**, *20*, 1354−1370.
(13) Soulika, A. M.; Morikis, D.; Sarrias, M. R.; Roy, M.; Spruce, L. A.; Sahu, A.; Lambris, J. D. *J. Immunol.* **2003**, *171*, 1881−1890.
(14) (a) Klepeis, J. L.; Floudas, C. A.; Morikis, D.; Tsokos, C. G.; Argyropoulos, E.; Spruce, L.; Lambris, J. D. *J. Am. Chem. Soc.* **2003**, *125*, 8422−8423. (b) Klepeis, J. L.; Floudas, C. A.; Morikis, D.; Tsokos, C. G.; Lambris, J. D. *Ind. Eng. Chem. Res.* **2004**, *43*, 3817−3826.
(15) (a) Katragadda, M.; Morikis, D.; Lambris, J. D. *J. Biol. Chem.* **2004**, *279*, 54987−54995. (b) Sahu, A.; Morikis, D.; and Lambris, J. D. *Mol. Immunol.* **2003**, *39*, 557−566.
(16) Mallik, B.; Lambris, J. D.; Morikis, D. *Proteins: Struct., Funct., Genet.* **2003**, *53*, 130−141.
(17) (a) Huang, P.; Kim, S.; Loew, G. J. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 21−28. (b) Brandt, W. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 615−621.
(18) (a) Zheng, Q.; Kyle, D. J. *Drug Discovery Today* **1997**, *2*, 229−234. (b) Walters, W. P.; Stahl, M. T.; Murco, M. A. *Drug Discovery Today* **1998**, *3*, 160−178.
(19) (a) Carlson, H. A.; Masukawa, K. M.; Rubins, K.; Bushman, F. D.; Jorgensen, W. L.; Lins, R. D.; Briggs, J. M.; McCammon, J. A. *J. Med. Chem.* **2000**, *43*, 2100−2114. (b) Carlson, H. A.; McCammon, J. A. *Mol. Pharmacol.* **2000**, *57*, 213−218. (c) Mustata, G. I.; Briggs, J. M. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 935−953. (d) Krovat, E. M.; Langer, T. *J. Med. Chem.* **2003**, *46*, 716−726.
(20) Bernard, D.; Coop, A.; MacKerell, D., Jr. *J. Am. Chem. Soc.* **2003**, *125*, 3101−3107.

Table 1. Nine Compstatin Analogues Used for the Pharmacophore Development[a]

	peptide	sequence	RIA[b]
		Active Analogues	
1	compstatin	I1-[C2-V3-V4-Q5-D6-W7-G8-H9-H10-R11-C12]-T13-NH_2	1
2	Ac-compstatin	**Ac**-I1-[C2-V3-V4-Q5-D6-W7-G8-H9-H10-R11-C12]-T13-NH_2	3
3	Ac-H9A	**Ac**-I1-[C2-V3-V4-Q5-D6-W7-G8-**A9**-H10-R11-C12]-T13-NH_2	4
4	Ac-V4Y/H9A	**Ac**-I1-[C2-V3-**Y4**-Q5-D6-W7-G8-**A9**-H10-R11-C12]-T13-NH_2	14
5	Ac-V4W/H9A	**Ac**-I1-[C2-V3-**W4**-Q5-D6-W7-G8-**A9**-H10-R11-C12]-T13-NH_2	45
		Inactive Analogues	
6	Ac-linear	**Ac**-I1-***C2***-V3-V4-Q5-D6-W7-G8-H9-H10-R11-***C12***-T13-NH_2	0
7	Ac-G8A	**Ac**-I1-[C2-V3-V4-Q5-D6-W7-**A8**-H9-H10-R11-C12]-T13-NH_2	0
8	Ac-Q5A	**Ac**-I1-[C2-V3-V4-**A5**-D6-W7-G8-H9-H10-R11-C12]-T13-NH_2	0
9	Ac-Q5G	**Ac**-I1-[C2-V3-V4-**G5**-D6-W7-G8-H9-H10-R11-C12]-T13-NH_2	0

[a] Acetyl group (Ac) and residues in bold-faced type represent differences from parent peptide compstatin. Italicized Cys residues represent a reduced disulfide bond in the linear analogue, called Ac-linear. [b] Relative inhibitory activities.[7−11,13,14]

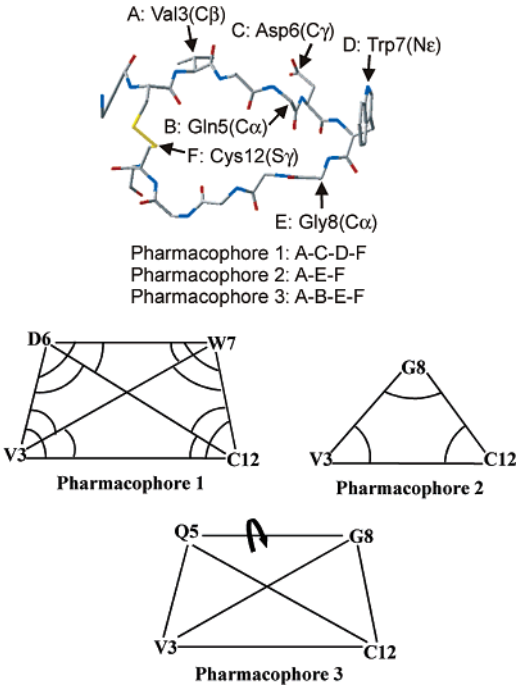


Figure 1. Topology of the selected pharmacophore points using a minimized structure of compstatin. Only side chains with pharmacophore points are shown, whereas the remaining side chains are deleted for clarity. The selected side chains, Cys2, Val3, Asp6, Trp7, and Cys12, contain physicochemical property pharmacophore points. The selected backbone C_α atoms are structural pharmacophore points. Two inactive analogues have Ala or Gly at position 5, and another has Ala at position 8 (Table 1). The terminal blocking groups are not shown. Three pharmacophore models considered in our study are also shown with all possible distances, angles, and a dihedral angle for each model separately.

side chain of Val3, and the side chains of the amino acids of the middle residues of the β-turn, Asp6 and Trp7. The remaining set involves the disulfide-bonded Cys2−Cys12, the side chain of Val3, and the backbone of the terminal amino acids of the β-turn at positions 5 and 8. The Cys2−Cys12 disulfide bond and Val3, both part of the hydrophobic surface patch, are indispensable for activity.[4,8] In addition, the β-turn, part of the polar surface, is essential for activity.[4,8] The pharmacophore points were carefully chosen and consist of structural or physicochemical property elements which were important for the structural stability and activity of compstatin. The physicochemical/structural properties per pharmacophore point chosen are as follows: Val3, nonpolar/coil; Gln5, polar/β-turn; Asp6, polar/β-turn; Trp7, nonpolar/β-turn; Gly8, polar/β-turn; Cys12, nonpolar/disulfide bond. All of them are possible hydrogen bond donors or acceptors too (because of nearby NH or CO groups). Specifically, during the optimization of compstatin, the seven amino acids Cys2, Val3, Gln5, Asp6, Trp7, Gly8, and Cys12 were deemed necessary for activity and were kept fixed, while the remaining six amino acids were allowed to vary.[4,11,13,14]

The Pharmacophore 1 is a four-point model, with selected set of pharmacophore points the C_β atom of Val3, the C_γ atom of Asp6, the $N_{\epsilon 1}$ atom of Trp7, and the S_γ atom of Cys12 for the Cys2−Cys12 disulfide bond. The pharmacophore 2 is a three-point model, with selected set of pharmacophore points the C_β atom of Val3, the C_α atom of Gly8 (or Ala8 for inactive analogue Ac-G8A), and the S_γ atom of Cys12 for the Cys2−Cys12 disulfide bond. The Pharmacophore 3 is also a four-point model, with selected set of pharmacophore points the C_β atom of Val3, the C_α atom of Gln5 (Ala5 or Gly5 for inactive analogues Ac-Q5A or Ac-Q5G, respectively), the C_α atom of Gly8 (Ala8 for inactive analogue Ac-G8A), and the S_γ atom of Cys12 for the Cys2−Cys12 disulfide bond. Figure 1 shows graphically the selection and geometry of the pharmacophore points. Pharmacophores 1 and 2 are studied in detail taking into account all combinations of interpoint distances and angles. The Pharmacophore 3, in addition to all distances, takes into account the effect of the dihedral angle of atoms Val3(C_β)−Residue 5(C_α)−Residue 8(C_α)−Cys12(S_γ), where Residue 5 is Gln, Ala, Gly and Residue 8 is Gly, Ala (Table 1).

The NMR solution structure of compstatin consists of a family of 21 low-energy structures.[8] In our earlier MD simulations of compstatin we considered these 21 structures plus an average minimized structure and a global optimization structure as the input.[16] Out of these 23 structures, Model 8 was the lowest-energy structure. In the present study, we have thus chosen Model 8 to represent compstatin. For the remaining compstatin analogues we have generated theoretical mutants using Model 8 of the NMR structure of compstatin. All of the analogues used (Table 1) contain the acetyl-blocking group ($-COCH_3$) at the N-terminus and the amide blocking group ($-NH_2$) at the C-terminus. Theoretical mutants have been constructed by using CHARMM with toph19/param19 polar-hydrogen topologies and parameters.[21] All-hydrogen topology and parameters for the N-terminal acetyl group have been taken from toph22/param22 set and adapted in toph19/param19 set.

Compstatin and its eight analogues (Table 1) were subjected to 300 steps of adopted-basis Newton Raphson (ABNR) gradient minimization before MD to relax the possible atomic contacts. A 5 ns MD simulation using the Verlet algorithm implemented in CHARMM (version c28b1) was performed for each of the nine structures at room temperature. SHAKE has been applied to all covalent bonds involving hydrogen.[22] Integration time step was 2 fs. Infinite cutoff distances were on for all nonbonded interactions with constant dielectric of 80. Both minimization and simulations were performed with implicit solvent effects using the generalized Born (GB) continuum solvent model.[23] The MD trajectories were saved every 10 ps. To identify the minimum-energy conformers we have sampled the MD trajectory every 50 ps for

(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) Ryckaert, J.-P.; Cicotti, G.; Berendsen, H. C. *J. Comput. Phys.* **1977**, 23, 327−341.
(23) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127−6129

Table 2. Distances and Angles among Pharmacophore Points in Active Analogues Using the Minimized Lowest-Energy Structures (Mini) and All Structures (All) during the MD Trajectories[a]

calculated parameter	compstatin		Ac-comps		Ac-H9A		Ac-V4Y/H9A		Ac-V4W/H9A	
	mini[b]	all[c]	mini[b]	all[c]	mini[b]	all[c]	mini[b]	all[c]	mini[b]	all[c]
					Distance					
V3−D6	7.5	7.9	13.2	12.3	8.7	8.8	10.3	11.5	11.5	11.2
V3−W7	10.2	10.5	10.8	11.7	9.7	9.5	9.9	11.6	10.4	11.9
V3−C12	8.4	8.1	8.3	8.8	8.2	8.5	7.8	8.0	8.2	8.2
D6−W7	5.7	5.5	7.6	7.7	6.0	6.5	5.6	7.7	5.3	5.4
D6−C12	14.5	14.5	16.3	16.7	14.6	14.3	16.0	17.1	16.6	16.8
W7−C12	15.6	15.0	13.7	12.5	11.8	11.0	12.9	13.4	13.1	14.5
V3−G8	7.8	8.0	9.5	11.0	4.4	4.6	11.5	10.8	7.0	7.0
G8−C12	10.4	10.2	10.1	12.2	7.2	7.0	14.2	12.8	13.6	13.3
					Angle					
V3−D6−W7	100.1	102.2	55.5	61.3	79.9	76.6	69.7	76.4	64.6	80.4
D6−W7−V3	46.6	43.7	89.3	80.3	62.5	64.9	78.2	65.1	88.1	67.9
W7−V3−D6	33.3	32.4	35.2	35.5	37.5	40.6	32.1	39.9	27.3	27.2
V3−D6−C12	25.4	26.3	30.4	29.2	29.4	29.1	24.0	25.0	27.2	29.9
D6−C12−V3	22.6	24.4	53.1	45.2	31.5	32.9	32.6	34.8	39.8	35.3
C12−V3−D6	132.0	129.7	96.5	104.3	119.1	120.0	123.4	120.6	112.9	115.1
V3−W7−C12	29.5	31.5	37.4	44.8	43.7	48.1	37.4	36.7	39.0	29.7
W7−C12−V3	36.8	40.2	52.4	56.6	54.7	57.7	50.6	56.5	52.7	51.1
C12−V3−W7	113.7	106.3	90.2	75.5	81.6	78.3	91.9	86.6	88.3	101.6
D6−W7−C12	68.8	70.1	96.1	80.7	106.3	104.7	115.0	99.2	123.0	97.6
W7−C12−D6	21.4	22.5	27.5	27.4	23.2	25.0	18.5	26.2	15.5	18.9
C12−D6−W7	89.8	82.4	56.4	61.4	50.5	46.8	46.5	52.7	41.5	66.6
V3−G8−C12	52.6	53.1	50.2	42.7	86.4	88.3	33.2	34.6	29.7	30.4
G8−C12−V3	47.6	49.7	61.3	60.0	32.6	35.2	53.8	53.5	25.0	22.7
C12−V3−G8	79.8	75.6	68.5	78.0	61.0	56.5	93.1	94.4	125.4	128.1
					Dihedral Angle					
V3−Q5−G8−C12	72.9	67.4	65.2	42.9	118.1	148.2	36.7	50.0	17.3	24.0

[a] Distances are given in Å and angles in degrees. [b] Values correspond to the lowest-energy structure identified from the MD trajectory (see Methods). [c] Represents highest probability value from the entire MD trajectory.

efficiency. This sampling generated 100 structures, each one of which was subsequently subjected to 300 steps of ABNR energy minimization. Out of the 100 minimized structures the one with lowest energy was selected.

One-dimensional probability distributions of distances, angles, and dihedral angles and two-dimensional probability distributions of distance−angle and distance−dihedral angle were calculated for the 500 conformers from snapshots of the MD trajectories. Distance parameters, minimum and maximum for a specified distance between two pharmacophoric points, were identified from the 500 conformers. Then, distance probabilities were calculated using a 0.5 Å distance bin. The number of distances within each bin divided by the total number (500 in the present study) represented the probability. Because of the inherent flexibility of the peptides in solution, we have tested a number of lower and higher bin widths. The best value of 0.5 Å was chosen for all analogues because, for lower bin widths, probabilities remained very low in the entire distance range that did not clearly indicate a maximum. Similarly, the probability distributions for angles and dihedral angles among pharmacophore points were calculated with an optimum angle bin of 5°.

Overall, for any three-point combination there are three distances and three angles, which produce nine permutations of distance−angle pairs or nine distance−angle contour plots. For any four-point combination there are six distances and 12 angles, which produce 36 distinct permutations of distance−angle pairs or 36 distance−angle contour plots. Therefore, for the four-point Pharmacophore 1, there are 36 contour plots and for the three-point Pharmacophore 2, there are nine contour plots for a single analogue. In total, there are 45 contour plots for Pharmacophores 1 and 2, giving rise to 405 contour plots for all nine active and inactive analogues. In addition, for the four-point Pharmacophore 3, there are six sets of distance−dihedral angle permutations or six contour plots. Therefore, for all three pharmacophores there are 51 contour plots (36 for Pharmacophore 1 + 9 for Pharmacophore 2 + 6 for Pharmacophore 3) representing 459 individual contour plots in total for all nine analogues.

Results

Three pharmacophore models have been constructed (Figure 1) using the structure of compstatin and structure−activity relations from previous rational design, which was part of a search for higher activity analogues. The four-point Pharmacophore 1, Val3(C_β)−Asp6(C_γ)−Trp7($N_{\epsilon 1}$)−Cys12(S_γ), is a mixed representation of important physicochemical (polar/nonpolar, charge, hydrogen bonding) characteristics. The three-point Pharmacophore 2, Val3(C_β)−Residue 8(C_α)−Cys12(S_γ), incorporates a structural characteristic through the backbone C_α atom at position 8. In the case of active analogues Gly8 is the last (fourth) residue of the β-turn (Table 1; Figure 1). The four-point Pharmacophore 3, Val3(C_β)−Residue 5(C_α)−Residue 8(C_α)−Cys12(S_γ), incorporates two structural characteristics through the backbone C_α atoms at positions 5 and 8, which in the case of active analogues are the first and last residue of the β-turn (Gln5 and Gly8; Table 1; Figure 1). One-dimensional probability distributions for every distance and angle in Pharmacophores 1 and 2 and for the C_β−C_α−C_α−S_γ dihedral angle in Pharmacophore 3 were calculated for active and inactive analogues (Tables 2 and 3; Figures 2, 3, 4, and SI). To support our choice of pharmacophore points, pairwise residue−residue interaction energies have been calculated for all analogues and are shown in Figure 5. Two-dimensional probability distributions for all possible combinations of distance−angle pairs in Pharmacophore 1 and Pharmacophore 2 (Figure 6A−F and SI) and distance−dihedral angle pairs in Pharmacophore 3 (Figure 6G−L) were also calculated.

Table 2 shows a list of all distances, angles, and dihedral angle of the three pharmacophores for the active compstatin analogues. The listed values correspond to the minimized lowest-

Table 3. Distances and Angles among Pharmacophore Points in Inactive Analogues Using the Minimized Lowest-Energy Structures (Mini) and All Structures (All) during the MD Trajectories[a]

calculated parameter	Ac-linear		Ac-G8A		Ac-Q5A		Ac-Q5G	
	mini[b]	all[c]	mini[b]	all[c]	mini[b]	all[c]	mini[b]	all[c]
				Distance				
V3−D6	6.7	7.2	5.7	7.2, 12.2	12.0	11.3	5.5	4.7
V3−W7	11.7	12.0	10.4	10.6	12.0	11.2	10.0	10.4
V3−C12	9.6	9.3	6.9	7.4	8.9	9.1	4.6	4.8,7.3
D6−W7	8.6	8.6	6.7	5.9	6.0	7.1	4.7	7.5
D6−C12	13.7	13.9	11.2	13.9	12.6	13.4	5.6	5.2
W7−C12	11.8	11.3	13.5	13.5	9.5	17.1	9.0	9.4
V3−G8[c]	10.9	10.9	8.3	8.5	6.6	9.8	10.0	10.1
G8−C12[c]	12.7	12.3	12.8	11.4	10.8	11.4	12.1	12.0
				Angle				
V3−D6−W7	99.2	98.3	108.3	99.6	75.5	72.6	156.2	125.7
D6−W7−V3	34.3	39.6	32.6	39.5	75.5	74.0	12.8	19.6
W7−V3−D6	46.4	46.8	39.1	31.2	29.0	35.9	11.0	29.2
V3−D6−C12	40.7	39.8	30.1	22.2	42.3	41.6	49.3	55.7
D6−C12−V3	26.9	34.0	24.6	27.7	65.3	56.7	63.8	56.4
C12−V3−D6	112.5	104.5	125.3	131.1	72.3	86.1	66.9	42.7,62.7
V3−W7−C12	48.5	48.6	29.6	25.9	47.1	27.9	27.6	28.5
W7−C12−V3	65.4	68.2	46.3	49.2	81.4	32.5	88.8	97.5
C12−V3−W7	66.1	61.2	104.0	106.4	51.4	120.3	63.7	46.0
D6−W7−C12	82.8	84.9	55.8	65.3	106.3	50.4	33.0	35.3
W7−C12−D6	38.6	35.1	29.6	26.5	27.3	25.0	27.4	36.9
C12−D6−W7	58.6	52.1	94.6	82.9	46.4	111.6	119.6	98.5
V3−G8−C12	47.5	48.5	29.1	33.7	55.2	45.9	21.4	21.7
G8−C12−V3	56.7	58.1	36.0	46.0	37.8	43.0	51.9	63.4
C12−V3−G8	75.8	73.9	114.9	104.0	87.0	77.6	106.7	69.6
				Dihedral Angle				
V3−Q5−G8−C12	106.1	112.8	48.1	54.0	52.5	51.0	32.3	30.2

[a] Distances are given in Å and angles in degrees. [b] Values correspond to the lowest-energy structure identified from the MD trajectory (see Methods). [c] Represents highest probability value from the entire MD trajectory.

energy structure and the highest one-dimensional probability from all structures of the MD trajectory. Table 3 shows a list of all distances, angles, and dihedral angle of the three pharmacophores for the inactive compstatin analogues. In total, for all three pharmacophores, there are 8 distances, 15 angles, and one dihedral angle (Tables 2 and 3). For both active and inactive analogues, the most probable distances are very close to their corresponding values in the lowest-energy minimized conformers. This means that the major conformers are well represented by the lowest-energy structure in the conformational energy landscape.

Figure 2 shows representative plots of one-dimensional probability distributions for V3−D6, W7−C12, and V3−G8 distances in active and inactive compstatin analogues. Figure 3 shows representative plots of one-dimensional probability distributions for V3−D6−W7, D6−W7−C12, and V3−G8−C12 angles in active and inactive compstatin analogues. The complete set of plots is deposited as SI. Figure 4 shows plots of the one-dimensional probability distributions for V3−Residue 5−Residue 8−C12 dihedral angle. There is a great variety in the shape (height, width, and number of maxima) in the calculated one-dimensional probability distributions. In most cases, there is also overlap in the one-dimensional probability distributions of active or inactive compstatin analogues. In general, plots of the type of Figures 2−4 (and of SI) do not allow for the discrimination of the physicochemical and conformational preferences between active and inactive compstatin analogues, using our selected pharmacophore points. For this reason we plotted the same data in the form of two-dimensional probabilities for (distance, angle) and (distance, dihedral angle) pairs.

Since the structural parameters in maximum probability conformers and minimum energy conformers are in the same range (Tables 2 and 3), we calculated pairwise van der Waals (VDW) and electrostatic energies for all nine analogues using the minimum-energy conformers. The plots are presented in Figure 5. A few pairwise interactions potentially important for activity are distinguished in these plots. The highest-activity analogue Ac-V4W/H9A (Table 1) possesses a distinct favorable VDW interaction between Trp4 and Trp7, which was previously attributed to $\pi-\pi$ interaction between the rings of the two tryptophans.[11] The two tryptophan rings in this analogue are oriented almost parallel to each other, reflecting this favorable interaction. In other active analogues, consistently a favorable VDW interaction between Gln5 and Trp7 is observed, which is absent only in Ac-V4W/H9A. This suggests that in inactive analogues, absence of Gln5 disturbs the VDW interaction

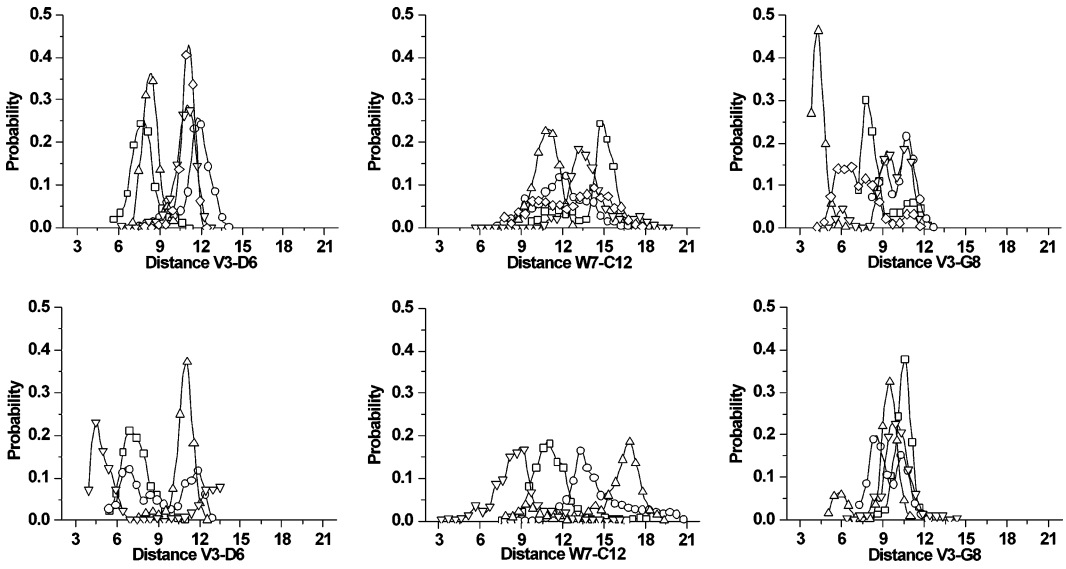


Figure 2. One-dimensional probability distributions for selected distances between pharmacophore points. The upper panels correspond to active analogues, and the lower panels correspond to inactive analogues (Table 1). The active analogues are compstatin (squares), Ac-compstatin (circles), Ac-H9A (triangles), Ac-V4Y/H9A (inverted triangles), and Ac-V4W/H9A (diamonds). The inactive analogues are Ac-linear (squares), Ac-G8A (circles), Ac-Q5A (triangles), Ac-Q5G (inverted triangles).

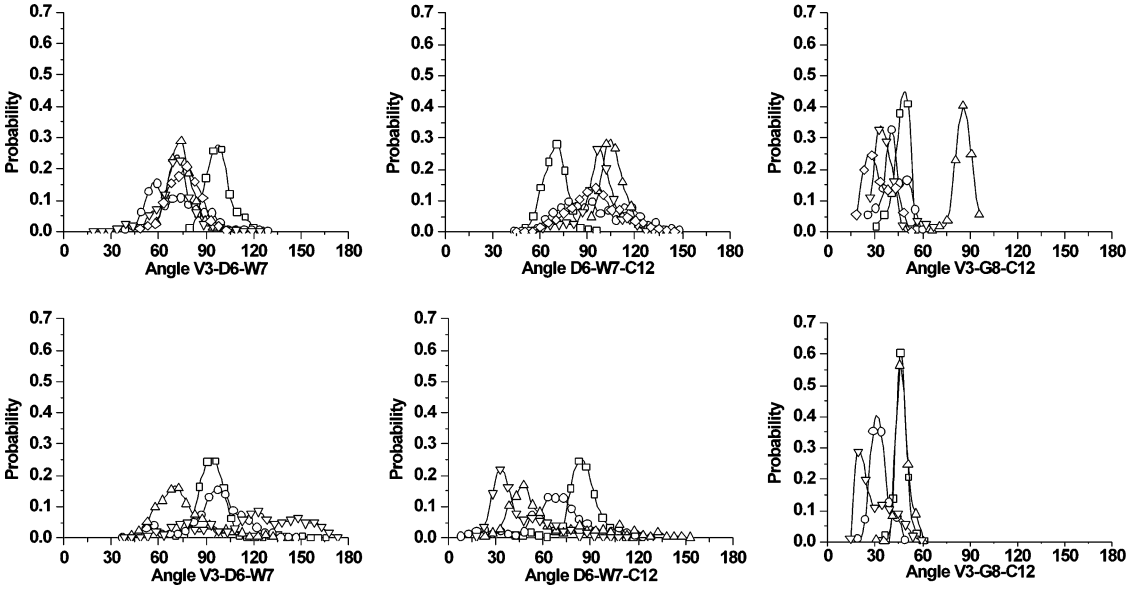


Figure 3. One-dimensional probability distributions for selected angles between pharmacophore points. The upper panels correspond to active analogues, and the lower panels correspond to inactive analogues (Table 1). The active analogues are compstatin (squares), Ac-compstatin (circles), Ac-H9A (triangles), Ac-V4Y/H9A (inverted triangles), and Ac-V4W/H9A (diamonds). The inactive analogues are Ac-linear (squares), Ac-G8A (circles), Ac-Q5A (triangles), Ac-Q5G (inverted triangles).

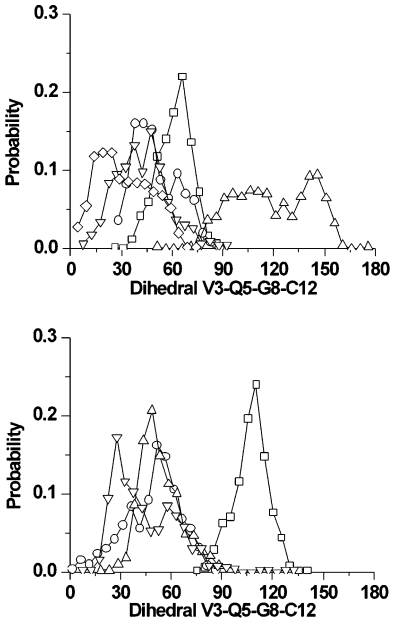
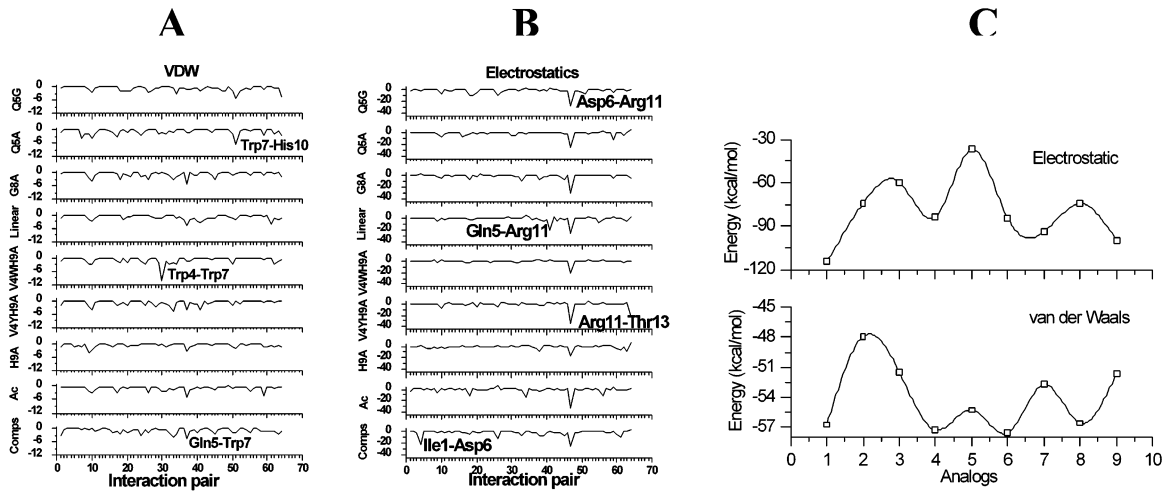


Figure 4. One-dimensional probability distributions for the dihedral angles between pharmacophore points. The upper panel corresponds to active analogues, and the lower panel corresponds to inactive analogues (Table 1). The active analogues are compstatin (squares), Ac-compstatin (circles), Ac-H9A (triangles), Ac-V4Y/H9A (inverted triangles), and Ac-V4W/H9A (diamonds). The inactive analogues are Ac-linear (squares), Ac-G8A (circles), Ac-Q5A (triangles), Ac-Q5G (inverted triangles).

causing the ring of Trp7 to drift away. In the analogues Ac-Q5A and Ac-Q5G, a strong favorable VDW interaction involving Trp7 and His10 is observed. The proper orientation of Trp7 is thus one of the primary factors important for the activity. In the linear compstatin analogue a −25 kcal/mol electrostatic stabilization is gained by the Gln5-Arg11 interaction alone. Because this analogue does not contain disulfide bond, reorientation of side chains near the two termini makes it possible for Gln5 to come in closer proximity to Arg11. However, Gln5-Arg11 electrostatic interaction might not be important for activity. The residue Asp6 is part of the β-turn in the active analogues. Although the β-turn is disrupted in the inactive analogues, the orientation of Asp6 side chain with respect to Arg11 is consistently similar in all analogues. This implies that the possibility of strong electrostatic interaction between opposite charges or even the possibility of formation of a salt bridge barely influences the activity pattern. Few other electrostatic interactions, such as Ile1-Asp6 in compstatin and Arg11-Thr13 in Ac-V4Y/H9A, are quite strong but do not appear to participate in ligand−receptor interaction directly. The rest of the pairwise VDW and electrostatic interactions are not noticeably different in all of the analogues. This indicates that more factors in addition to the topology of individual residues in the free analogues are responsible for discriminating active and inactive analogues. Also, this supports our choice not to include residues at positions 1, 4, 9, 10, 11, and 13.

Two-dimensional probability contour plots for all distance−angle permutations (Pharmacophore 1 and Pharmacophore 2) have been generated with the aim to discriminate active and inactive analogues. Plots A−F of Figure 6 show representative two-dimensional probability contour plots, and the complete set of plots can be found in SI. The relative spatial positions of pharmacophore points Val3(C_β), Asp6(C_γ), Trp7($N_{\epsilon 1}$), and Cys12(S_γ) in Pharmacophore 1 generate a small number of distance−angle two-dimensional contours that separate the analogues only partially in terms of their activity. The (distance, angle) contour plots (V3−D6, V3−D6−W7), (V3−D6, V3−D6−C12), and (W7−C12, W7−C12−V3) create a trend that clusters together Ac-V4Y/H9A, Ac-V4W/H9A, and Ac-Q5A and either compstatin or Ac-compstatin (Figure 6A−C). The remaining analogues can be seen as forming the second cluster. The (distance, angle) contour centers for the various analogues are better separated in Figure 6D−F. The (distance, angle) plots (D6−W7, D6−W7−C12) and (D6−C12, D6−W7−C12) more effectively cluster active and inactive analogues (Figure 6D,E).



1: 1-3	2: 1-4	3: 1-5	4: 1-6	5: 1-7	6: 1-8	7: 1-9	8: 1-10	9: 1-12	10: 1-13
11: 2-4	12: 2-5	13: 2-6	14: 2-7	15: 2-8	16: 2-9	17: 2-10	18: 2-11	19: 2-13	20: 3-5
21: 3-6	22: 3-7	23: 3-8	24: 3-9	25: 3-10	26: 3-11	27: 3-12	28: 3-13	29: 4-6	30: 4-7
31: 4-8	32: 4-9	33: 4-10	34: 4-11	35: 4-12	36: 4-13	37: 5-7	38: 5-8	39: 5-9	40: 5-10
41: 5-11	42: 5-12	43: 5-13	44: 6-8	45: 6-9	46: 6-10	47: 6-11	48: 6-12	49: 6-13	50: 7-9
51: 7-10	52: 7-11	53: 7-12	54: 7-13	55: 8-10	56: 8-11	57: 8-12	58: 8-13	59: 9-11	60: 9-12
61: 9-13	62: 10-12	63: 10-13	64: 11-13						

Figure 5. Pairwise (A) van der Waals and (B) electrostatic energies (in kcal/mol) of the five active and four inactive compstatin analogues (Table 1). (C) Sum of the pairwise interaction energies. Analogues are (1) compstatin, (2) Ac-compstatin, (3) Ac-H9A, (4) Ac-V4Y/H9A, (5) Ac-V4W/H9A, (6) Ac-linear, (7) Ac-G8A, (8) Ac-Q5A, (9) Ac-Q5G. (D) Key to the interaction pairs (horizontal axis) for (A) and (B). The energies were calculated using minimum-energy structure of each analogue (see text).

The use of structural backbone pharmacophoric point in Pharmacophore 2 did not help in discriminating active from inactive analogues (Figure 6F and SI). The C_α atom of residue Gly8 was chosen because Gly8 is the last residue of the β-turn and was previously found to be indispensable for structural stability and activity, which are correlated.[10] Specifically, the lack of a side chain in Gly was deemed necessary to increase the available conformational space, allowing the β-turn to be formed.[8]

Because the discrimination between active and inactive analogues using the two-dimensional probabilities was inconclusive, we extended our study to examine if a strategically selected dihedral angle produces a better discrimination between activity and inactivity. This dihedral angle is defined by pharmacophore points Val3(C_β)−Residue 5(C_α)−Residue 8(C_α)−Cys12(S_γ), where Residue 5 is Gln (for active analogues), Ala, or Gly and Residue 8 is Gly (for active analogues) or Ala. This is our Pharmacophore 3 model (Figure 1). The one-dimensional probability distributions of dihedral angle are given in Figure 4. The probability maxima of individual analogues are distinct, but large overlaps between 0° and 75° are observed for most of the analogues. However, the Ac-H9A and linear analogues have entirely different region of preference (Figure 4). No conclusive activity correlations can be drawn from these data, which led us consider two-dimensional probabilities for distance−dihedral pairs.

The distance−dihedral angle contour plots for Pharmacophore 3 are shown in Figure 6G−L. The integrity of the β-turn is depicted in Figure 6G, which shows the plot of the dihedral angle against the distance of the C_α atoms of the end-residues of the β-turn, Gln5-Gly8. The dihedral angle also reflects the torsion around the line that connects the C_α atoms of the end-residues of the β-turn. A β-turn is possible when the $C_\alpha-C_\alpha$ distance of its end-residues is less than 7 Å.[24] The β-turn structure is clearly lost in the inactive analogue Ac-Q5A because its $C_\alpha-C_\alpha$ distance is greater than 8 Å (Figure 6G), but the remaining analogues comply with this end-residue criterion (Figure 6G). The inactive linear analogue is separated from the cyclic analogues in all plots (Figure 6G−L). The linear analogue has a propensity for the formation of a turn of a 3_{10} helix.[10] A turn of a 3_{10} helix is similar to a type-III β-turn, which is considered as a variation of a type-I β-turn. The active Ac-H9A analogue is also separated from the rest (Figure 6G−L). The incorporation of Ala at position 9 increases the peptide propensity for helical conformation.[10,11] The contour plots involving the distances Q5−C12 and V3−G8, the diagonals of the skew quadrilateral V3−Q5−G8−C12, discriminate active from inactive analogues with acceptable resolution (Figure 6H,K). The contour plots of Figure 6H,I,K separate the most active analogues, Ac-V4W/H9A and Ac-V4Y/H9A, from the rest (Figure 6H,K), depicting the effect of aromaticity at position 4 on activity.[11] For the highest-activity analogue Ac-V4W/H9A, the V3−G8 distance is rather short (about 6 Å; Figure 6K). Together with the small dihedral angle value (about 20°), this indicates that in the analogue Ac-V4W/H9A, Val3 is favorably placed between the two strands and closer to the β-turn while Trp4 is favorably placed in proximity to the turn residue Trp7. The V3−G8 distance in the active analogues compstatin, Ac-compstatin, and Ac-V4Y/H9A is much longer, and the dihedral angle is slightly larger than in the Ac-V4W/H9A analogue (Figure 6K). It is possible that, within similar β-turn structure, the residues at positions 3 and 4 act cooperatively to increase

(24) Schulz, G. E.; Schirmer, R. H. *Principles of Protein Structure*; Springer-Verlag: New York, New York, 1979.

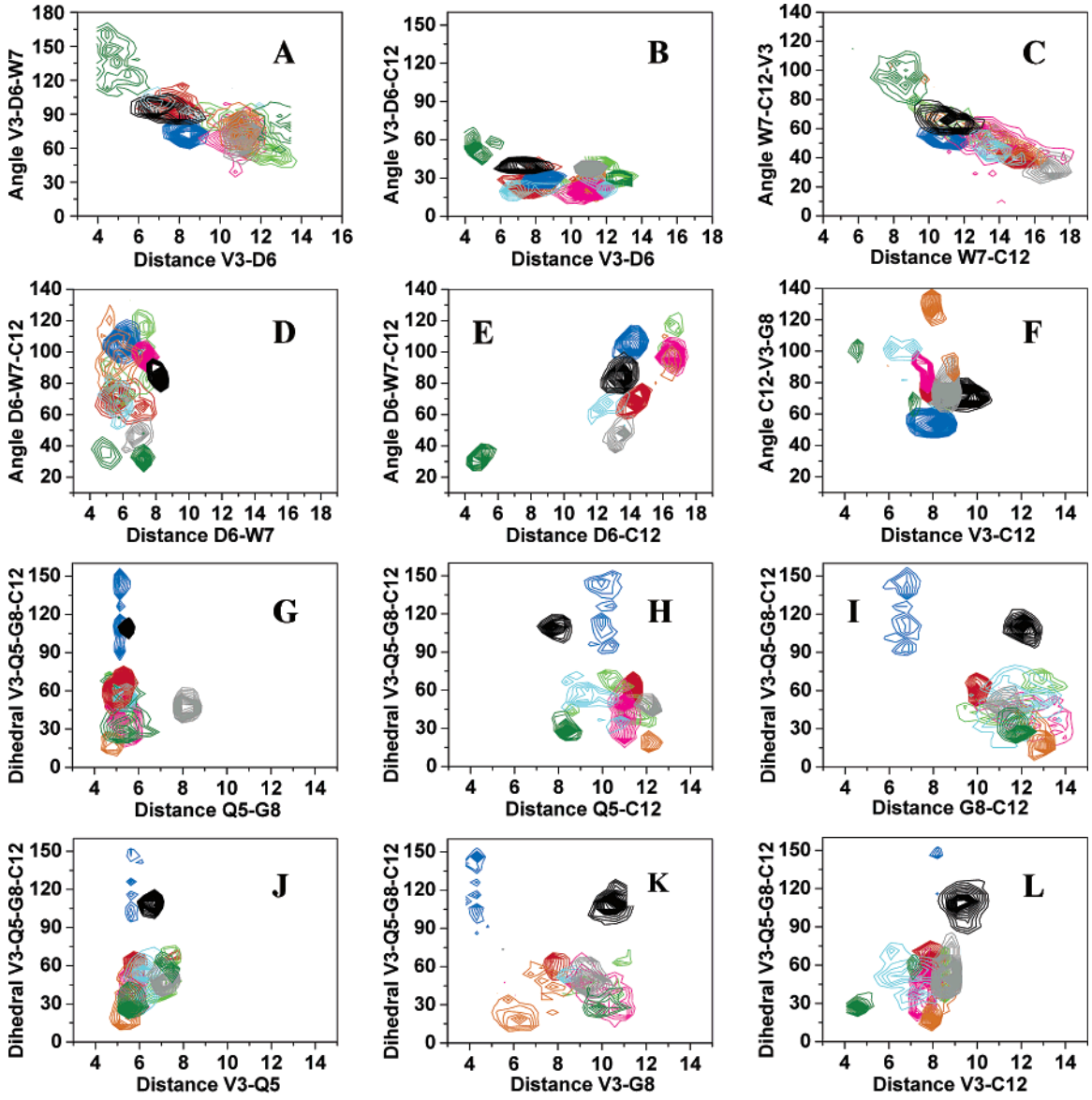


Figure 6. Two-dimensional probability distribution contours for (A−F) selected distance−angle pairs from Pharmacophore 1 and Pharmacophore 2 and (G−L) distance−dihedral angle pair. The contour interval was fixed to 0.005 units in all plots. The contour color code is red for compstatin, green for Ac-compstatin, blue for Ac-H9A, magenta for Ac-V4Y/H9A, orange for Ac-V4W/H9A, black for Ac-linear, cyan for Ac-G8A, gray for Ac-Q5A, olive (dark green) for Ac-Q5G.

the activity. The remaining active analogue, Ac-H9A, has distinct behavior (shorter distance but larger dihedral angle; Figure 6K) and is discussed above.

In summary, the choice of our dihedral angle, V3-Q5-G8-C12 (Pharmacophore 3; Figure 1), together with the two diagonal distances, V3−G8 and Q5−C12, provide an efficient way to discriminate active from inactive analogues using two-dimensional probability distributions. In addition, the contour plots involving the dihedral angle are capable of depicting structural subtleties important for activity. Overall, our observations reinforce the important roles of the β-turn,[8] the disulfide bridge,[10] the hydrophobic cluster[10] and Trp4−Trp7 interaction[11] for activity.

Discussion

We have presented the construction of three quasi-dynamic pharmacophore models from first principles, using active and inactive analogues of the 13-residue cyclic peptide compstatin. Our goal was to develop pharmacophore models for the spatial arrangement of the physicochemical properties of compstatin activity alone or in combination with structural properties. The underlying assumption is that the selected physicochemical and structural properties are the most important features of free compstatin and its analogues for activity. The selection of the physicochemical and structural properties was based on previous structure-based optimization of the activity of compstatin using rational and combinatorial approaches.[4,8−11,13,14] Tools for structure determination were (i) NMR spectroscopy in combination with computational methods, such as MD-based simulated annealing[8] or global optimization,[12] (ii) MD simulations for an NMR ensemble of structures,[16] and (iii) alanine scan or site-specific amino acid replacements.[8,10] The combinatorial approaches were either experimental at the DNA level using phage-displayed random peptide libraries[13] or computational using global optimization methods.[12,14] Activities were measured

using standard immunological assays for classical or alternative pathway complement activation as reported.[6] In the present study we have performed MD simulations on the structures of the parent peptide compstatin and four active and four inactive compstatin analogues to generate three related but distinct pharmacophore models. The initial structures for the MD simulations were the NMR-derived structure of compstatin[8] or models of the remaining analogues based on the NMR-derived structure of compstatin and theoretical amino acid replacements. Structural snapshots during the MD trajectory provided the structural templates for our pharmacophore models. The term "quasi-dynamic" refers to the selection of structural snapshots at specific time intervals from the dynamic trajectory to construct otherwise static pharmacophore models.

Using this knowledge we have selected one three-point and two four-point pharmacophore models. One of the four-point models contains exclusively side-chain atoms to represent pharmacophore points. The three-point and the other four-point models have mixed side-chain and backbone atoms. Probability distributions of all distances and angles for Pharmacophore 1 and Pharmacophore 2 and only of a dihedral angle for Pharmacophore 3 are calculated. Two-dimensional contours of every distance−angle combination (Pharmacophore 1 and Pharmacophore 2) and distance−dihedral angle combination (Pharmacophore 3) along a conformational trajectory have been created. The trajectory contains 500 conformations saved at 10 ps intervals in a 5 ns MD simulation.

The goal of our study was to determine the best mixed physicochemical/structural property pharmacophore, which could separate active from inactive analogues or alternatively analogues of high activity from analogues of low activity. For this reason we have selected compstatin, four active analogues of compstatin, and four inactive analogues of compstatin as our training set. These pharmacophore models can be used to search nonpeptidic compound databases for possible active compounds, which will resemble the important physicochemical and structural characteristics of compstatin for binding and activity.

Our study was inspired by a paper by Bernard et al.,[20] who have used the dynamic conformers of nonpeptidic δ-opioid agonists and antagonists to develop a three-point pharmacophore model for the differentiation of these two classes of ligands. Their model was based on the probability distribution of mutual distances and angles among selected functional moieties from a large set of conformers generated by MD simulation. The functional moieties were considered to be important in receptor recognition. Expectedly, large overlaps of the individual probability distribution of distances and angles occurred among different analogues.[20] In our study, in addition to similar overlaps, many distances and angles cover a wide range, implying that our peptides are internally very flexible compared to the more rigid organic molecules. Bernard et al.,[20] selected six prototypical agonists and four antagonists to develop the model and one compound structurally similar to the agonists to validate the model. Because of the greater structural flexibility, they did not include available endogenous peptide ligands in their model building training set. Despite the greater rigidity in the compound skeleton and higher sampling frequency of their MD simulations (necessary because of rigidity), the probability distributions of individual distances and angles failed to discriminate the two classes of ligands. This observation, according to the authors, was not disappointing because their study indicated a direct relation between the selected physicochemical properties and pharmacophore points, despite the fact that there was no direct correlation with the overall geometries of the compounds.

We have adopted, tested, and extended the validity of the idea of Bernard et al.[20] for peptides as opposed to organic molecules. The training set for our pharmacophore models consists of the relatively large (13-residue) peptide compstatin and several analogues. To identify the structural parameters essential to discriminate analogues in terms of their activities and to translate them into a dynamic pharmacophore model, a large number of two-dimensional probability distribution contour plots are generated for distance−angle and distance−dihedral angle combinations. Although active and inactive analogues do not form distinct nonoverlapping clusters, individual analogues are identifiable in specific regions of the plots. In some plots there is good contour resolution along the distance (horizontal) axis, indicating preference in all analogues for stable angles or dihedral angles. In other plots there is good contour resolution along the angle or dihedral angle (vertical) axis, indicating preference in all analogues for a stable distance. There are also cases where there is good resolution throughout the whole plot areas, and these are most useful for analogue discrimination. In our training set, compstatin and seven out of its eight analogues are cyclic peptides. Although a disulfide bond and a β-turn impose sufficient rigidity in the backbone movement, side chains are free to move. Overall these peptides are much more flexible compared to the organic molecules with extensive networks of bulky rings. It may be that the inclusion of side chains, which are carriers of the amino acid physicochemical properties, are responsible for the observed great degree of contour overlap (Figure 6A−F). Bernard et al.[20] considered a three-point pharmacophore model, where two points were not the atoms, but they were centroids of either an aromatic ring or a hydrophobic region. The relative positions of the aromatic ring centroids were very similar in all molecules, but centroids of hydrophobic regions were different owing to variation of the hydrophobic moieties. The maximum range in any distance was about 2 Å, whereas the maximum angle range was no more than 30°. This was reflected in both the one- and two-dimensional contour plots. The maximum probability regions for the groups of agonists and antagonists, although obvious in most plots, were in close proximity. The distinction between the groups of agonists and antagonists possibly reflected the effect of structural rigidity and choice of pharmacophore points.[20] In our case distinction of active and inactive analogues based on probability distribution plots of distances, angles, or distance−angle pairs is not clear (Figures 2, 3, 6A−F). Nonetheless, we have extended our analysis to include probability distribution plots of a dihedral angle and distance−dihedral angle pairs, aiming to reduce the overlaps between active and inactive analogues (Figures 4, 6G−L). The choice of including the dihedral angle in the two-dimensional probability distribution plots revealed preferred trends for distance−dihedral angle pairs and improved the separation of the two groups of active and inactive analogues. For example, the optimum distance criterion between the two end-residues of the β-turn has been established by the (Q5−G8, V3−Q5−G8−C12) contour plot (Figure 6G). Also, the two diagonal distances, V3−

G8 and Q5−C12, coupled with the dihedral angle, V3−Q5−G8−C12, are capable of discriminating the active from the inactive analogues (Figure 6H,K). In addition, the highest-activity analogue is discriminated from the remaining active analogues and the inactive analogues.

Quantitative estimations of residue−residue pairwise nonbonded interaction energies have indicated that every individual analogue in the training set shows a very similar pattern, which does not allow the discrimination of active from inactive analogues. For a set of structurally similar and flexible peptides such as compstatin and its analogues, the use of dynamic conformers does not necessarily establish quantitative structure−activity relations in terms of structural parameters such as distance, angle, and dihedral angle. Neither the structures of the major conformers nor the nonbonded energies of dynamically sampled conformers of our individual peptides exclusively group them as active or inactive. The aim of our study was to build a pharmacophore model by identifying substructures in terms of selected geometric parameters and physicochemical properties. The final model in the present case is the spatial arrangement, using distances, angles, and dihedral angles, of a combination of substructures involving nonpolar groups, polar groups, β-turn, disulfide bridge, hydrogen bond donors and acceptors, hydrophobic centers, etc. The ideal geometric parameters for constructing this model are those that separate active and inactive analogues.

Conclusions

On the basis of a wealth of structural and physicochemical properties of the 13-residue peptide compstatin and a large number of its analogues, we have selected the parent peptide and four active and four inactive analogues as a training set for the construction of three quasi-dynamic pharmacophore models from first principles. Our goal was to discriminate active from inactive analogues, based on their structure and dynamics. Sampling of the conformers along MD trajectories and calculation of one-dimensional probability distributions of several important distances, angles, and dihedral angles and two-dimensional distance−angle and distance−dihedral angle pairs were the keys to the construction of these models. Comparison of our study with a similar study previously applied to a set of more rigid organic molecules[20] indicates that larger and more flexible molecules such as peptides require higher-order structural parameters, beyond simple distance and angle, to discriminate the sets of active and inactive analogues. While analysis of two-dimensional probability distribution contour plots involving a large set of distance−angle pairs failed, a limited number of distance−dihedral angle pairs provided a way to discriminate active from inactive analogues. In addition, the probability distribution contour plots involving the dihedral angle allowed for the depiction of structural subtleties important for activity. The original idea of Bernard et al.[20] for conformationally sampled pharmacophores has been validated and extended here, using a more flexible peptide training set. We expect that our extension using distance−dihedral angle pairs will be useful to build pharmacophore models for organic molecules as well. In closing, two-dimensional probability distribution analysis involving dynamic conformers can be followed as an efficient route for the ligand-based pharmacophore development of both peptidic and nonpeptidic ligands. We expect that the success of this method will vary for different peptide systems and for different selections of the type and number of pharmacophore points. The success of this method depends on the number of active and inactive analogues in the training set and on how well the molecular dynamics protocol samples the conformational space. Flexibility is a common aspect of peptide structure and function and should be taken into consideration in ligand-based pharmacophore design and ligand binding studies in general.

Acknowledgment. This work was supported by NIH Grant R24 GM069736.

Supporting Information Available: (1) Figures of one-dimensional probability distributions of distances, angles, and dihedral angles; (2) figures of two-dimensional distance−angle and distance−dihedral angle probability distribution contour plots; (3) tables of highest probability distance−angle and distance−dihedral angle pairs. This material is available free of charge via the Internet at http://pubs.acs.org.

JA051004C