Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 8506–8510, November 1990
Biochemistry

Elucidating the structural chemistry of glycosaminoglycan recognition by protein C inhibitor

(molecular modeling/electrostatic complementarity/serpins/heparin/amphipathic α-helices)

LESLIE A. KUHN*†, JOHN H. GRIFFIN†‡, CINDY L. FISHER*, JUDITH S. GREENGARD†, BONNO N. BOUMA§, FRANCISCO ESPAÑA†¶, AND JOHN A. TAINER*

Departments of *Molecular Biology and †Molecular and Experimental Medicine and the Committee on Vascular Biology, Research Institute of Scripps Clinic, La Jolla, CA 92037; and §Department of Haematology, University Hospital, Utrecht, The Netherlands

Communicated by Elkan Blout, July 2, 1990

ABSTRACT **Glycosaminoglycans (GAGs) including heparin accelerate the inhibition of serine proteases by serine protease inhibitors (serpins), an essential process in regulating blood coagulation. To analyze the molecular basis for GAG recognition by the plasma serpin protein C inhibitor (PCI; also known as plasminogen activator inhibitor 3), we have constructed a complete, energy-minimized, three-dimensional model of PCI by using the structure of homologous α_1-antitrypsin as a template. Sequence analysis, hydrogen-bonding environment, and shape complementarity suggested that the N-terminal residues of PCI, which are not homologous to those of α_1-antitrypsin, form an amphipathic α-helix, here designated A+ since it precedes the α_1-antitrypsin A helix. Electrostatic calculations revealed a single, highly positive surface region arising from both the A+ and H helices, suggesting that this two-helix motif is required for GAG binding by PCI. The dominant role of electrostatic interactions in PCI–heparin binding was confirmed by the strong ionic strength dependence of heparin stimulation. The involvement of the A+ helix in heparin binding was verified by demonstrating that an anti-PCI antibody that specifically binds the A+ peptide blocks heparin binding.**

Noncovalent associations of glycosaminoglycans (GAGs) with proteins mediate a variety of biological processes, including cell attachment, growth, and differentiation (1) and anticoagulation (2). Although the binding of GAGs to proteins has important biological consequences, the structural chemistry of this binding has not yet been determined. Structural and chemical features conserved among GAG-binding proteins may provide insights into the binding process. Protein C inhibitor (PCI) (3, 4) inhibits activated protein C (APC) in an interaction stimulated by heparin, and PCI has overall sequence homology to two serine protease inhibitors (serpins) that have been studied extensively: antithrombin III (ATIII) and α_1-antitrypsin (α_1AT). Biochemical data on ATIII have identified several positive residues involved in heparin binding, but these residues, mainly found in the D helix of the schematic ATIII structural model (5), are not positively charged in PCI. To determine whether PCI and ATIII share a common structural basis for heparin binding and to allow molecular surface and electrostatic calculations, we have constructed a complete, energy-minimized, three-dimensional model of PCI based on its homology to α_1AT. Examination of surface shapes and electrostatic potentials identified a single, highly positive area as the probable heparin binding site. Experimental results verified the dominance of electrostatic interactions between heparin and PCI and showed that an antibody binding the N-terminal peptide eliminated heparin binding.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

Construction of an Energy-Minimized Model of PCI. PCI was modeled based on its sequence homology to α_1AT, for which a high-resolution crystallographic structure has been determined (6). The sequence of human PCI (4) was aligned with the human α_1AT sequence (7) by two methods: (*i*) manually, using the criterion that chemically similar (i.e., hydrophobic, hydrophilic, positively charged, negatively charged) residues be optimally aligned without introducing gaps in either sequence and (*ii*) automatically, using the program ALIGN (8) with both mutation data and identity matrices, with gap penalties ranging from 2 to 10. ALIGN results were not used because many gaps were introduced that only marginally improved the sequence homology, even when a large gap penalty was used. The structure of PCI residues 20–391 (α_1AT numbering) was built by side-chain substitution with the molecular editor MOLEDT (Biosym Technologies, San Diego), using the x-ray structure of α_1AT (Protein Data Bank entry 6API) as a template and following the original side-chain torsion angles. Side-chain collisions were corrected by using new torsion angles from a rotamer library (9), and small adjustments were made where necessary by using the program INSIGHT (Biosym Technologies). α_1AT and side-chain-substituted PCI structures were then energy minimized by using the program DISCOVER (Biosym Technologies) with an all-atom force field. Steepest descent minimization was applied while forcing heavy atom (non-H) positions to a template of the starting model, decreasing the force constant stepwise from 1000 to 10 kcal·Å^{-1} (1 cal = 4.184 J) until the potential energy derivative was <10 kcal·mol^{-1}·Å^{-1} for each atom. The energies were further minimized by using the conjugate gradient algorithm with a force constant of 10 kcal·Å^{-1} applied only to main-chain atoms until the maximum potential energy derivative was <10 kcal·mol^{-1}·Å^{-1}. This methodology is similar to a recently published homology modeling method (10).

Since the 15 N-terminal residues (residues 5–19) of PCI are not sequence homologous to α_1AT and the N terminus of α_1AT is not visible in the electron density map (6), the structure of these residues in PCI was modeled independently. α-Helical, β-strand, and loop models were built with MOLEDT and were analyzed in several orientations on the surface of the minimized PCI model by using INSIGHT. To

Abbreviations: PCI, protein C inhibitor; GAG, glycosaminoglycan; APC, activated protein C; ATIII, antithrombin III; α_1AT, α_1-antitrypsin.

‡To whom reprint requests should be addressed.
¶Present address: Hospital La Fe, Centro Investigacion, 46009 Valencia, Spain.

evaluate shape complementarity, the solvent-accessible molecular surfaces of the N terminus and PCI models were calculated by using the program MS (11) with a 1.6-Å probe sphere and surface dot density of 2 dots per Å^2. For the most plausible orientations of the N terminus on the surface of PCI, the surface buried between the N terminus and PCI was calculated with MS and was analyzed by using the graphics programs GRAMPS, GRANNY, and INSIGHT. These N-terminus models were then evaluated for hydrogen bonding and charge complementarity with the PCI surface. For the best conformations and orientations of the N-terminal residues, the peptide bond between residue 19 (the C terminus of the N-terminal segment) and residue 20 (the N terminus of residues 20–391 of PCI) was made by optimally orienting the N-terminal segment and then making minimal changes in the backbone torsion angles of residues 19 and 20 and their nearest neighbors. Plausible models of PCI including the N terminus were energy minimized as described above to alleviate unfavorable residue contacts and to improve the conformations of the turn at residues 19 and 20 and the N-terminal segment.

Quantification of the Electrostatic Surface Potential of PCI. Coulombic electrostatic potentials were calculated at all points on the solvent-accessible surfaces of the energy-minimized PCI models by using the programs ESPOT and ESSURF (12) with two different dielectric models: the dielectric constant of bulk water, $\varepsilon = 80$, and a linear distance-dependent dielectric, $\varepsilon = 4r$. The electrostatic potential value for each molecular surface point was color coded.

The role of electrostatics in the interaction of heparin with PCI was assessed experimentally by measuring the ionic strength influence on heparin stimulation of the formation of PCI–APC complexes. The second-order rate constant, k_2, for the inhibition of APC by PCI was determined as described (13). PCI (2–1600 nM) in 0.01 M Tris·HCl/1% bovine serum albumin/0.02% NaN_3, pH 7.4, with NaCl added to the indicated ionic strengths, was incubated for 5 min at 37°C with heparin at the indicated concentrations, after which APC (2 nM) was added. At timed intervals, aliquots were diluted 1:30 with 1 mM chromogenic substrate S-2366 (Kabi Vitrum, Stockholm) in 0.05 M Tris·HCl/0.1 M NaCl/4 mM $CaCl_2$/0.02% NaN_3, pH 8.2, to determine residual APC activity.

Analysis of Heparin Binding by Using Anti-PCI Antibodies. The specificity of murine monoclonal antibodies raised against intact purified PCI (14), as determined by ELISA and peptide competition assays using component peptides of PCI, will be published elsewhere as part of a comprehensive study (J.S.G., B.N.B., and J.H.G., unpublished data). Two anti-PCI antibodies characterized by the above study were used to test PCI binding in this paper: API39, shown to bind the A+ helix peptide, and API60, which acted as a control. IgG was purified by protein A-Sepharose chromatography (Pharmacia) as recommended by the manufacturer. PCI (4 μg) was incubated for 60 min at 22°C with API39 (48 μg), API60 (48 μg), or buffer, in 200 μl of 0.01 M Tris·HCl/0.14 M NaCl, pH 7.4. The sample was adjusted to 0.1 M NaCl in a final vol of 400 μl and loaded onto a 0.6 ml-column of heparin-agarose (Sigma). Using a Pharmacia FPLC system, PCI was eluted (0.1-ml fractions) with a linear gradient from 0.1 to 0.6 M NaCl. The elution profiles of PCI antigen were determined by an ELISA (15).

RESULTS AND DISCUSSION

Comparison of PCI and α_1AT Sequences. In the absence of high-resolution structural data, the construction of a three-dimensional model for a protein based on its homology to a known structure permits structure-guided experimentation and the rational design of therapeutic agents. For this study, we used the structure of human α_1AT (6) as a template for three-dimensional modeling of the homologous serpin PCI. The alignment of human PCI (4) and human α_1AT (7) sequences was based on manual alignment of chemically similar residues and involved no insertions or deletions (Fig. 1). The high degree and uniformity of the sequence homology between PCI and α_1AT suggested that the structure of PCI is largely similar to that of α_1AT. Of the residues of PCI modeled on α_1AT (20–391 in the α_1AT numbering system), 71.8% were conserved within hydrophobicity categories (16) and 44.4% were identical. The two proteins were homologous along the entire length of PCI except for the 27 N-terminal residues and residues 351–382, which surround the reactive center cleavage site at residue 358. Realignment of this segment would have involved introducing two multiresidue gaps without significantly improving the number of identical residues in the alignment. A study of the relationship between atomic coordinates in modeled and actual structures of a number of proteins (17) has shown that the expected rms deviation between C_α positions in proteins having ≈45% sequence identity is 1.0 Å. Moreover, the serpin fold seems to be highly conserved even at a significantly lower level of sequence homology; the recent x-ray structure of plakalbumin has a main-chain atom rms deviation of <1 Å from the α_1AT structure, although the two sequences are only 30% similar (18).

Models of the N Terminus of PCI. Analysis of the α-helix, β-strand, and loop models of the 15 N-terminal residues showed that an α-helical structure was most likely. Possible β-strand and loop structures were deemed implausible be-

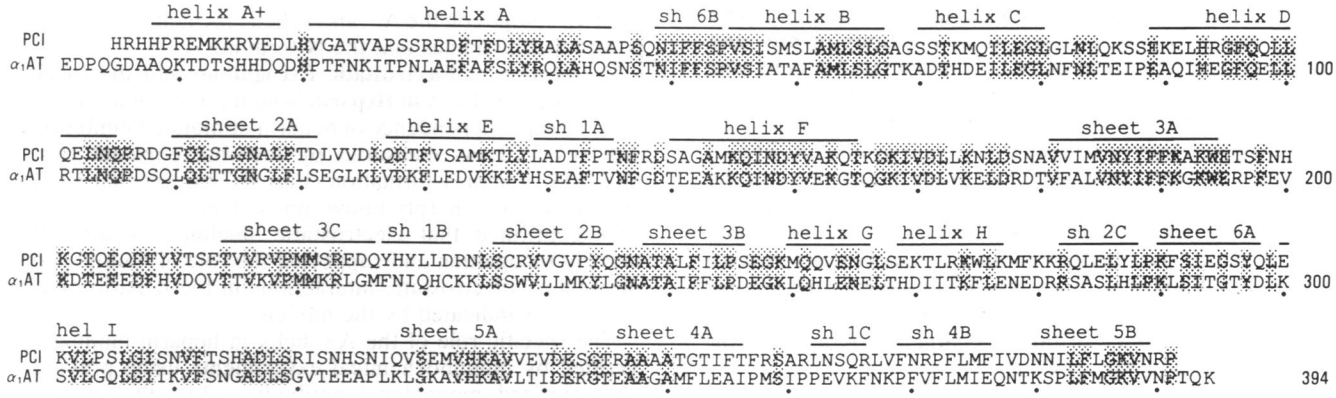

FIG. 1. Sequence alignment of human PCI with human α_1AT. Residues are numbered (with dots under every 10th residue) and helices and sheets are named as in α_1AT. Assignment of secondary structure elements in PCI was based on the minimized PCI structure (see Fig. 3*b*). Identical residues are shown by stippling.

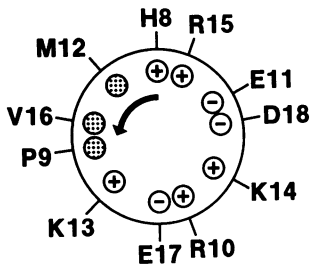

FIG. 2. α-Helical projection of residues 8–18 of PCI, the proposed amphipathic A+ helix. Stippled circles represent hydrophobic residues.

cause of their length, the lack of favorable side-chain interactions, and the absence of suitable main-chain hydrogen bonds. The helicity of these residues in the PCI model was further supported by two secondary structure algorithms (19, 20), which both predicted that the first nine residues of PCI (residues 9–17 in α_1AT numbering) fold as an α-helix. When represented on an α-helical wheel (Fig. 2), residues 8–18 formed a highly amphipathic structure with only hydrophobic residues on one side and only charged residues on the other. Studies of known protein structures show that sequence amphipathicity correlates with structural periodicity, and in particular that the sequences of surface α-helices are often amphipathic with a period of ≈3.6 residues (21).

The N-terminal region formed by residues 5–19, designated the A+ helix, was found to be complementary in shape to two surface grooves extending from residue 20, and in each groove the α-helix directed its charged side chains toward the solvent while burying its hydrophobic side chains against hydrophobic residues of PCI. Detailed analyses were made of the two structures resulting from energy minimization of PCI modeled with the A+ helix attached in both orientations: alongside the A helix at an angle of 25° to the H helix (model I) and antiparallel to the H helix at an angle of 40° (model II).

Evaluation of the Complete PCI Model. Changes in heavy atom positions during energy minimization of α_1AT and PCI were moderate and distributed throughout the proteins' structures. The rms deviation between heavy atom positions before and after energy minimization was 1.16 Å for PCI without the A+ helix (0.84 Å for C_α) and 1.17 Å for α_1AT (0.77 Å for C_α). Superposition of the energy-minimized structure of PCI without the 15 N-terminal residues on that of α_1AT showed their close similarity (Fig. 3*a*), with an overall rms deviation of 0.83 Å between postminimization C_α positions in PCI and α_1AT (residues 20–391). The most significant deviations in C_α positions between PCI and α_1AT were in the H helix and apparently resulted from the large number of repulsive positive charges in this area of PCI. The rms deviation between heavy atom positions in pre- and postminimized structures of PCI including the A+ helix was 1.21 Å (0.87 Å for C_α) for model I and 1.15 Å (0.83 Å for C_α) for model II. After minimization, the N terminus of the A+ helix was less tightly coiled, a common feature in known structures. Minimization also improved the geometry of the turn at residues 19 and 20, with the rest of the PCI structure being substantially the same as that of PCI minimized without the A+ helix. The energy-minimized models I and II were judged reasonable based on the moderate rms deviations, conservation of hydrophobic core residues and secondary structure, the absence of buried charged residues, and the appropriateness of turn residues.[‖]

[‖]PCI models I and II have been deposited in the Protein Data Bank, Chemistry Department, Brookhaven National Laboratory, Upton NY 11973 (PDB code 1PAI).

Characterization of the Region of PCI Responsible for Electrostatic Recognition of Heparin. To identify positive regions of PCI that could bind the negatively charged groups of GAGs, the electrostatic potential at the surface of PCI was analyzed. PCI exhibited an overall electrostatic dipole, with a highly positive region including the H helix opposed by a weakly negative region centered on Asp-121 (Fig. 3*a*). In the electrostatic potential surfaces calculated with a dielectric constant of 80, the models including the A+ helix had a single, highly positive (≥3 kcal·mol^{-1}) surface region centered on Arg-10 and Lys-274, which protruded from the A+ and H helices (Fig. 3 *b* and *c*). Other central positive residues were Lys-14 and Lys-270 in model I (Fig. 3*b*) and Arg-6, Lys-277, and Lys-280 in model II (Fig. 3*c*). The positive region in both models formed a single face of the protein, had an area (1365 Å^2 in model I and 1705 Å^2 in model II) consistent with other protein interfaces (22), and was large enough to bind an extended heparin octasaccharide chain. Positive residues contributing at least 30 Å^2 to the ≥3 kcal·mol^{-1} surface in model I were residues 5–8, 10, 13–15, 86, 270, 273–274, and 277 and in model II were residues 6–7, 10, 13, 270, 273–274, 277, and 280–282. All of these residues except Lys-86 were either in the A+ or H helices or immediately followed the H helix. The region with potential ≥3 kcal·mol^{-1} constituted 9% of the total surface of PCI in model I and 11% in model II.

Electrostatic calculations done on PCI surfaces either without the A+ helix or with the H helix charges neutralized (data not shown) indicated that neither the A+ helix nor the H helix alone generated a strongly positive surface; in both cases, there was <19 Å^2 of surface with potential ≥3 kcal·mol^{-1}. There was also no significant positive surface associated with the D helix, which is implicated in heparin binding to ATIII (5). These results suggest that in PCI both helices A+ and H, but not D, are essential for GAG binding. The distance-dependent dielectric model gave qualitatively similar results.

The number and arrangement of charges on the A+ and H helices closely complement the charges of a heparin octasaccharide. Approximately eight extended, chair-form saccharide units, each ≈4 Å long (23), fit into the groove between the A+ and H helices in either PCI model I or II. This groove, ≈10 Å wide at its narrowest, could accommodate the common saccharide units of heparin (23), iduronic acid (≈7 Å wide), and N-sulfated D-glucosamine 6-sulfate (≈10 Å wide). Thus, the negatively charged saccharide rings could bridge the gap between the positively charged side chains of the bordering A+ and H helices, with the conformational flexibility inherent in glycosidic linkages and protein side chains allowing these interactions to be optimized. An octasaccharide from heparin contains 12–16 negative sulfate and carboxylate groups (2) that could interact favorably with the 12 positive charges on the A+ and H helices (including the three N-terminal histidines and the N terminus).

Verification of Electrostatic Recognition and the Involvement of the A+ Helix in Heparin Binding. Examination of the ionic strength dependence of heparin-stimulated inhibition of APC by PCI, as measured by the rate constant (k_2) of PCI–APC complex formation (Fig. 4), showed that heparin stimulation fell sharply above physiological ionic strength. This confirmed that electrostatic shielding impairs PCI–heparin recognition and supported the role of multiple complementary electrostatic interactions in the heparin binding process, as indicated by the models.

The specific role of the A+ helix in heparin binding was verified by immunochemical experiments using previously characterized monoclonal antibodies (14). PCI was prevented from binding to heparin-agarose (Fig. 5) by antibody API39, which recognizes the A+ helix amino acid sequence and neutralizes heparin stimulation of APC inhibition by PCI,

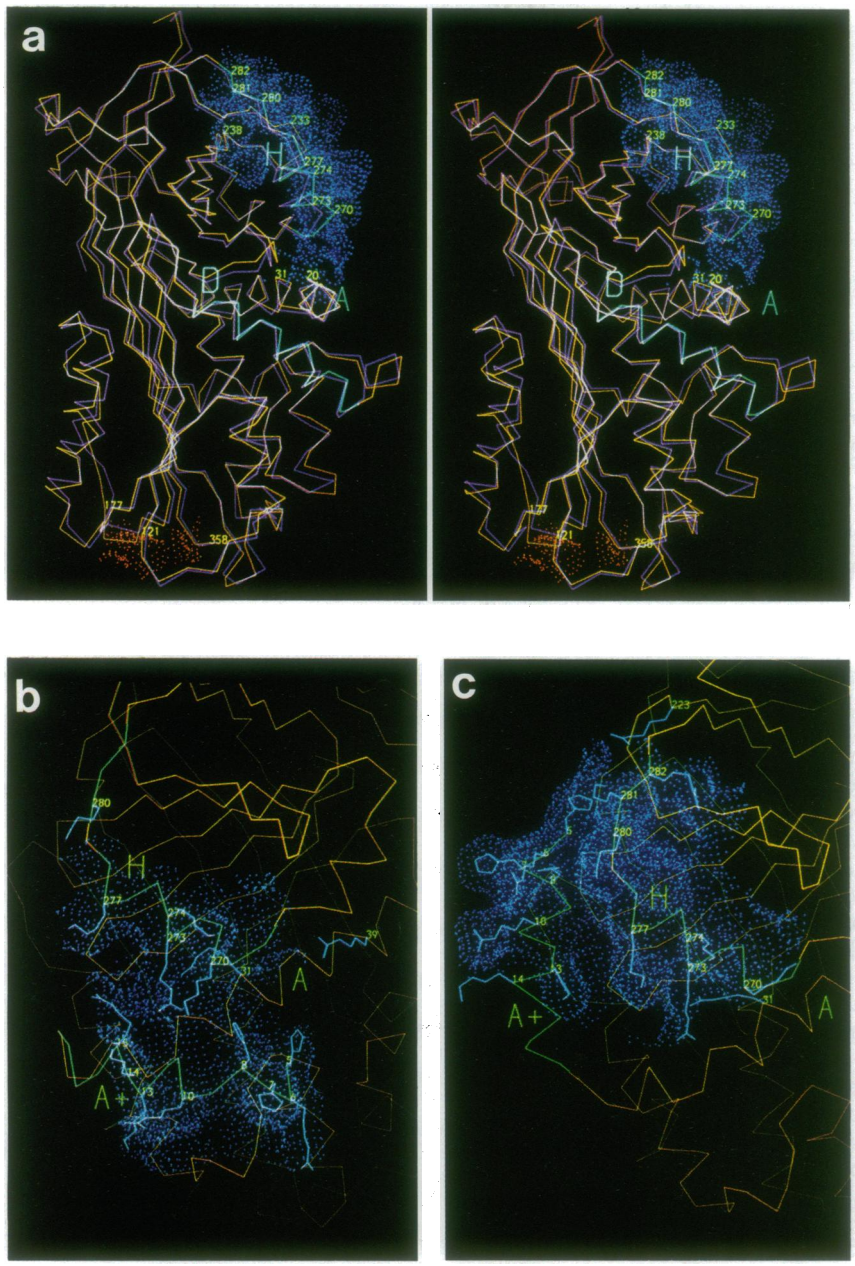

FIG. 3. (*a*) Stereo pair of the superimposed, energy-minimized C_α chains of residues 20–391 of α_1AT (magenta) and PCI (gold); also shown is the electrostatic potential surface of PCI without the N-terminal residues. The most positive (blue dots: potential ≥ 2 kcal·mol^{-1}) and most negative (red dots: potential ≤ -1 kcal·mol^{-1}) surface regions are displayed for PCI without the N-terminal residues. Labels (using α_1AT numbering) are given for charged residues contributing to the positive and negative surfaces. A, D, and H helices are indicated. (*b* and *c*) Orientation of the N-terminal residues and H helix and the positive electrostatic potential surface for energy-minimized structures of PCI model I (*b*) and PCI model II (*c*). Residues 8–18 of the N-terminal region form the A+ helix. The blue dot surface has an electrostatic potential of ≥ 3 kcal·mol^{-1}, which is 1 kcal·mol^{-1} more positive than the positive surface identified in *a*.

but not by antibody API60, which does not recognize either the A+ or the H helix amino acid sequences and does not affect heparin stimulation. These antibody binding experiments support the hypothesis that the A+ helix is in the heparin binding site of PCI.

Identification of a Two-Helix Motif for Heparin Recognition. The strikingly positive helix pair that forms the heparin recognition surface of PCI identified by this work is similar to the twin helical motif thought to bind heparin in dimers of platelet factor 4, a nonhomologous protein whose structure has recently been determined (24). In platelet factor 4, the two probable heparin-binding helices are antiparallel (related by a noncrystallographic twofold axis) and separated by 14 Å. Each helix is 12 residues long and has 4 lysine residues on its solvent-exposed side, and guanidination experiments suggest that these residues are responsible for heparin binding (25). In PCI, the A+ helix has 11 strictly helical residues including 2 lysines and 2 arginines on its solvent-exposed side. The PCI H helix has 9 residues including 3 solvent-exposed lysines and 1 arginine. In PCI model I, the A+ and H helices are separated by 9.4 Å at one end and 13.6 Å at the other end, while in model II, the helices are separated by 9.4 Å at one

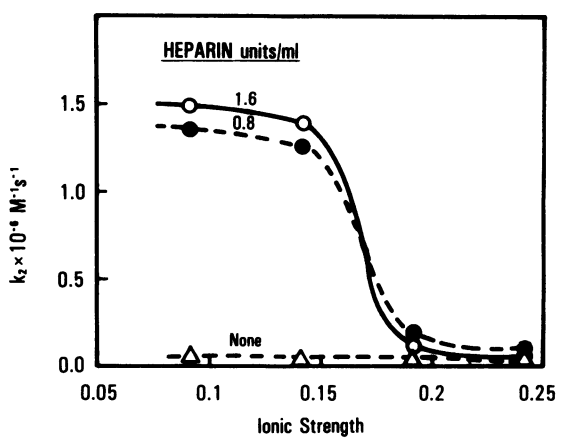

FIG. 4. Effect of ionic strength on heparin-dependent inhibition of APC by PCI. The second-order rate constant, k_2, for inhibition of APC by PCI was measured for a range of NaCl concentrations in the presence of heparin at optimum stimulatory concentrations and in the absence of heparin.

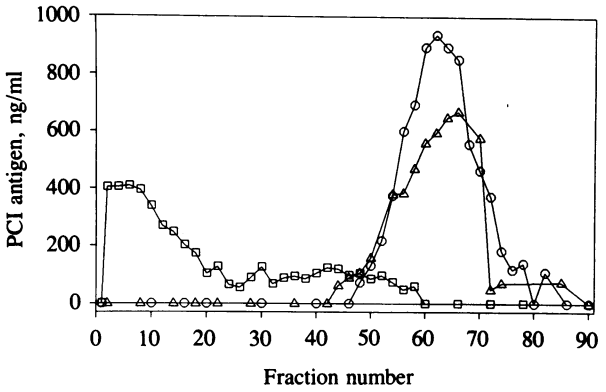

FIG. 5. Effect of murine monoclonal anti-PCI antibodies on binding of PCI to heparin-agarose. The elution profiles of PCI alone (○), PCI preincubated with API60 (△), and PCI preincubated with API39 (□) were determined by an ELISA for PCI antigen.

end and 21.9 Å at the other. Thus, the length of the helices, distance between helices, and the number of charges per helix are quite similar to those in platelet factor 4. In PCI, several positively charged groups close to the A+ and H helices (the N terminus; Arg-6; Lys-86; and His-5, -7, and -8, which are likely to be positively charged in the presence of heparin) further contribute to the region's positive potential and are likely to increase its affinity for heparin.

GAG recognition in ATIII may be a variation on this two-helix motif, involving positive residues in both the D helix (5) and a helix in the N-terminal region. Electrostatic surface calculations on ATIII modeled without the 44 N-terminal residues (L.A.K., J.A.T., and J.H.G., unpublished data) indicate that the D helix alone does not generate a large, positive surface, with only 64 $Å^2$ of the ATIII surface having an electrostatic potential ≥ 3 kcal·mol^{-1}. Thus, residues in the unmodeled N terminus of ATIII are likely required for electrostatic recognition of heparin and may provide an amphipathic helix that plays a role analogous to the A+ helix residues in PCI.

Sequential helix pairs in the helix-turn-helix DNA recognition motif are structurally conserved amongst DNA-binding proteins (26). For GAG-binding proteins, nonsequential amphipathic helix pairs may perform an analogous role, with the separation of the rigid, independent helices in serpins allowing conformational changes that affect the GAG binding function by optimizing shape and electrostatic complementarity. Beyond elucidating heparin–PCI interactions, the detailed structural model of PCI presented here will facilitate study of the interaction between APC and PCI and the design of inhibitor-resistant antithrombotic agents based on APC, as was recently done with tissue-type plasminogen activator (27).

We thank Elizabeth Getzoff, W. Bode, and R. Huber for valuable discussions. This research was supported by National Institutes of Health Grants GM39345 (J.A.T.) and HL31950 (J.H.G.), National Science Foundation Grant 8822385 (J.A.T.), and National Institutes of Health Training Grants DK07022 (L.A.K.) and GM11612 (C.L.F.).

1. Hassell, J. R., Kimura, J. H. & Hascall, V. C. (1986) *Annu. Rev. Biochem.* **55**, 539–567.
2. Casu, B., Petitou, M., Provasoli, M. & Sinaÿ, P. (1988) *Trends Biochem. Sci.* **13**, 221–225.
3. Marlar, R. A. & Griffin, J. H. (1980) *J. Clin. Invest.* **66**, 1186–1189.
4. Suzuki, K., Deyashiki, Y., Nishioka, J., Kurachi, K., Akira, M., Yamamoto, S. & Hashimoto, S. (1987) *J. Biol. Chem.* **262**, 611–616.
5. Carrell, R. W., Christey, P. B. & Boswell, D. R. (1987) in *Thrombosis and Haemostasis 1987*, eds. Verstraete, M., Vermylen, J., Lijnen, H. R. & Arnout, J. (Leuven University, Leuven, Belgium), pp. 1–15.
6. Loebermann, H., Tokuoka, R., Deisenhofer, J. & Huber, R. (1984) *J. Mol. Biol.* **177**, 531–556.
7. Carrell, R. W., Jeppsson, J.-O., Laurell, C.-B., Brennan, S. O., Owen, M. C., Vaughan, L. & Boswell, D. R. (1982) *Nature (London)* **298**, 329–334.
8. Dayhoff, M. O., Schwartz, R. M. & Orcutt, B. C. (1979) in *Atlas of Protein Sequence and Structure*, ed. Dayhoff, M. O. (Nat. Biomed. Res. Found., Washington, DC), Vol. 5, Suppl. 3, pp. 345–362.
9. Ponder, J. W. & Richards, F. M. (1987) *J. Mol. Biol.* **193**, 775–791.
10. Summers, N. L. & Karplus, M. (1989) *J. Mol. Biol.* **210**, 785–811.
11. Connolly, M. L. (1983) *Science* **221**, 709–713.
12. Getzoff, E. D., Tainer, J. A. & Olson, A. J. (1986) *Biophys. J.* **49**, 191–206.
13. España, F., Berrettini, M. & Griffin, J. H. (1989) *Thromb. Res.* **55**, 369–384.
14. Meijers, J. C. M., Vlooswijk, R. A. A., Kanters, D. H. A. J., Hessing, M. & Bouma, B. N. (1988) *Blood* **72**, 1401–1403.
15. España, F. & Griffin, J. H. (1989) *Thromb. Res.* **55**, 671–682.
16. Rose, G. D., Geselowitz, A. R., Lesser, G. J., Lee, R. H. & Zehfus, M. H. (1985) *Science* **229**, 834–838.
17. Blundell, T. L., Sibanda, B. L., Sternberg, M. J. E. & Thornton, J. M. (1987) *Nature (London)* **326**, 347–352.
18. Wright, H. T., Qian, H. X. & Huber, R. (1990) *J. Mol. Biol.* **213**, 513–528.
19. Garnier, J., Osguthorpe, D. J. & Robson, B. (1978) *J. Mol. Biol.* **120**, 97–120.
20. Chou, P. Y. & Fasman, G. D. (1978) *Adv. Enzymol. Relat. Areas Mol. Biol.* **47**, 45–148.
21. Eisenberg, D., Weiss, R. M. & Terwilliger, T. C. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 140–144.
22. Janin, J., Miller, S. & Chothia, C. (1988) *J. Mol. Biol.* **204**, 155–164.
23. Atkins, E. D. T. & Nieduszynski, I. A. (1976) in *Heparin: Chemistry and Clinical Usage*, eds. Kakkar, V. V. & Thomas, D. P. (Academic, London), pp. 21–35.
24. St. Charles, R., Walz, D. A. & Edwards, B. F. P. (1989) *J. Biol. Chem.* **264**, 2092–2099.
25. Handin, R. I. & Cohen, H. J. (1976) *J. Biol. Chem.* **251**, 4273–4282.
26. Richardson, J. S. & Richardson, D. C. (1988) *Proteins Struct. Funct. Genet.* **4**, 229–239.
27. Madison, E. L., Goldsmith, E. J., Gerard, R. D., Gething, M.-J. H. & Sambrook, J. F. (1989) *Nature (London)* **339**, 721–724.