Structure of Native Protein C Inhibitor Provides Insight into Its Multiple Functions*§

Wei Li‡, Ty E. Adams§, Margareta Kjellberg‡, Johan Stenflo§, and James A. Huntington†1

From the ‡Department of Haematology, Division of Structural Medicine, Thrombosis Research Unit, Cambridge Institute for Medical Research, University of Cambridge, Wellcome Trust/MRC Building, Hills Road, Cambridge CB2 2XY, United Kingdom and the §Department of Clinical Chemistry, University Hospital, Lund University, 5-20502 Malmö, Sweden

Protein C inhibitor (PCI) is a multifunctional serpin with wide ranging protease inhibitory functions, unique cofactor binding activities, and potential non-inhibitory functions akin to the hormone-transporting serpins. To gain insight into the molecular mechanisms utilized by PCI we developed a robust expression system in Escherichia coli and solved the crystal structure of PCI in its native state. The five monomers obtained from our two crystal forms provide an NMR-like ensemble revealing regions of inherent flexibility. The reactive center loop (RCL) of PCI is long and highly flexible with no evidence of hinge region incorporation into β-sheet A, as seen for other heparin-binding serpins. We adapted an extrinsic fluorescence method for determining dissociation constants for heparin and find that the N-terminal tail of PCI and residues adjacent to helix H are not involved in heparin binding. The minimal heparin length capable of tight binding to PCI was determined to be chains of eight monosaccharide units. A large hydrophobic pocket occupied by hydrophobic crystal contacts was found in an analogous position to the hormone-binding site in thyroxine-binding globulin. In conclusion, the data presented here provide important insights into the mechanisms by which PCI exercises its multiple inhibitory and non-inhibitory functions.

In many ways protein C inhibitor (PCI) (1) is a typical member of the serpin family of protease inhibitors (2, 3). It is a glycoprotein of 387 residues produced by several cell types and found in many tissues. In humans, PCI is produced in the liver and secreted into the blood where it circulates at ~100 nM (1).

PCI is named after its ability to inhibit the proteases which comprise the protein C anticoagulant pathway, activated protein C (APC) (4, 5) and the thrombin-thrombomodulin complex (6, 7), but it clearly inhibits different targets in tissues other than blood (8). However, identifying the true targets of PCI is made difficult by its inhibitory promiscuity. PCI utilizes the conserved serpin mechanism (9), where specificity is determined by kinetics of formation of the Michaelis complex (10). After which, the protease proceeds with attack of the scissile bond (P1–P1’, nomenclature of Schechter and Berger (12)), and at the acyl-enzyme intermediate stage the serpin undergoes a rapid conformational change which distorts the structure of the protease to prohibit deactivation (9, 14). This mechanism lends itself to promiscuous inhibition of proteases, and has been found to support the inhibition of subtilisin-type serine proteases (15) and some cysteine proteases (16, 17). Normally serpins have a single reactive center residue (P1) which significantly determines inhibitory specificity. An arginine in the P1 position of PCI makes it an inhibitor of trypsin-like proteases including uPA, tPA, thrombin, factors IXa, Xa, and XIa, plasmin, kallikrein, etc.; however PCI is also capable of inhibiting chymotrypsin-like proteases, presumably through the Phe at the P2 position (18). In some cases, exosites (residues which do not belong to the RCL) can also critically influence specificity (11). The general nature of the serpin mechanism combined with dual reactive centers renders PCI especially promiscuous, thus expanding the breadth of possible biological functions.

PCI belongs to a subclass of serpins known as the heparin-binding serpins, which includes antithrombin (AT), heparin cofactor II (HCII), protease nexin 1, and PAI-1 (for recent review see Ref. 19), but once again PCI stands out from the crowd. Whereas all other members utilize basic residues along the face of helix D for heparin binding, PCI is thought to utilize helix H and a basic N-terminal extension (20). The role of helix H has been confirmed by mutagenesis studies (21, 22), but the importance of the N-terminal region (‡HRHHPREMKKR‡) remains unresolved. It is, however, important to note that affinity of PCI for GAGs has never been measured directly, and that all inferences drawn concerning the importance of basic regions, GAG selectivity, or response to conformational change are drawn from elution from heparin-Sepharose or the effect on GAG acceleration of protease inhibition.

Another feature, which is commonly observed for heparin-binding serpins, is a heparin-induced conformational change (23). The crystal structures of native AT (24) and HCII (25) revealed the incorporation of the N terminus of the RCL (the
Crystal Structure of Native PCI

hinge region) into β-sheet A. Heparin activation is effected, in part, by liberating the RCL from its contacts with the body of the serpin (26, 27). Although PCI does not utilize helix D for heparin binding, a heparin-induced conformational change in PCI is possible, although it would presumably depend on the native structure of PCI resembling that of AT and HCII (i.e. hinge region inserted).

As for the other heparin-binding serpines, heparin acceleration of protease inhibition by PCI involves heparin binding to both the protease and PCI, thus implying a bridging mechanism (28). However, in contrast to other heparin-binding serpins, heparin binding to PCI can also decrease the rate of protease inhibition, as seen with tissue kallikrein (29). Thus, heparin-like GAGs can alter PCI specificity by either promoting or preventing protease inhibition, a unique feature that may play a key role in determining the tissue-specific activities of PCI.

The serpin architecture has also evolved to carry out some non-inhibitory functions such as hormone transport (30). Corticosteroid-binding globulin (CBG) and thyroxine-binding globulin (TBG) bind tightly to hydrophobic hormones and are thought to exploit the serpin conformational change for delivery to specific sites (e.g. sites of inflammation) (31, 32). PCI is also capable of interacting with hydrophobic molecules such as estradiol, testosterone, and retinoic acid (33), and although the functional relevance of this ability remains unclear, the mechanism of binding may be related to that utilized by CBG and TBG. Our structure of RCL-cleaved PCI revealed an unexpected shift in the position of helix H and a two-turn shortening of helix A, which created a large hydrophobic pocket, termed the helix A gap (34). The helix A gap was proposed to be the binding site for hydrophobic compounds, but it was not clear if this feature is preserved in native PCI.

Unlike many members of the serpin family, PCI has not been studied extensively at the biochemical and structural level, due in part, to the difficulty in obtaining large amounts of pure native protein. We have established an efficient method for the production of milligram quantities of wild-type and N-terminal truncated PCI from E. coli, which has allowed us to obtain diffraction quality crystals of native PCI under different conditions. Here we describe five native PCI structures obtained from two separate crystallization conditions, providing an opportunity to form an SDS-stable complex. We therefore left the tag in place for crystallization and biochemical characterization. The R229A mutation was generated according to the Stratagene site-directed mutagenesis protocol on the Δ16 background. Correct sequences for all clones were verified by DNA sequencing at Gene Service or Lark Technologies.

Expression and Purification of Recombinant PCI—PCI-pET19b vectors were transformed into Rosetta DE3 pLysS strain for protein expression. Briefly, cells were grown at 37 °C to mid-log phase followed by isopropyl-1-thio-β-D-galactopyranoside induction at 30 °C overnight. Purification was achieved with a Ni-NTA column (Amersham Biosciences), heparin-Sepharose, and S-75 gel filtration. We typically ran a final heparin Sepharose column after gel filtration in order to concentrate the protein, and to remove cleaved PCI, which eluted slightly later than native PCI. Typical yields of pure native PCI (wild-type and variants) were 2–3 mg per liter of culture.

Crystallization, Data Collection, and Refinement for C2 Crystal Form—Initial screens using both full-length and Δ16 PCI generated many leads. Diffraction-quality crystals were obtained for Δ16 PCI (2.6 mg/ml in 50 mM Tris-HCl pH 7.4, 550 mM NaCl) after seeding into drops containing 66.7 mM tripropylamine sulfate, 6.7% PEG3350, at room temperature. The crystal was cryoprotected in 12% PEG3350, 20% glycerol (increased in 5% steps), and 120 mM potassium citrate, and flash-cooled in vapor nitrogen to 100 K. Data were collected at Daresbury Synchrotron Radiation Source (Warrington, Cheshire, UK) station 14.1 from a single crystal, and processed using Mosflm, Scala,
Crystal Structure of Native PCI

TABLE 1
Data collection and refinement statistics

| Crystal form 1 | Crystal form 2 |
|---------------|---------------|
| **Crystal parameters** | | |
| Space group | C2 | P21 |
| Cell dimensions (Å) | | |
| a = 198.59, b = 70.95, c = 123.88, β = 127.92 | 51.8 | 55.8 |
| Solvent content (%) | | |
| | | |
| Data collection: | | |
| Wavelength (Å) | 1.488 | 0.931 |
| Resolution (Å) | 35-2.3 | 2.4-2.3 |
| Total reflections | 234342 | 29426 |
| Unique reflections | 60523 | 8725 |
| (I/σ(I)) | 13 | 2.9 |
| Multiplicity | 3.9 | 3.4 |
| Completeness (%) | 99.7 | 99.0 |
| Rmerge | 0.083 | 0.396 |
| Rmerge | 0.063 (0.062) | 0.217 (0.127) |

**Model details:**

| | |
|---|---|
| Atoms in asymmetric unit- | |
| Protein | 8256 | 5406 |
| Water | 437 | 284 |
| Glycerol molecules | 7 | 1 |
| Citrate ion | 1 | 0 |
| Average Bfactor (Å²) | 42.1 | 34.5 |
| Refinement statistics: | |
| Reflections (working set/free set) | 60214/2449 | 9434/400 |
| Rmerge (working set/free set) | 20.00-2.30 Å | 2.44-2.30 Å |
| Rmerge | 26.09-2.00 | 2.13-2.00 |
| Rmerge | 1857/104 | 27.0/30.5 |
| r.m.s.d. from ideal geometry- | |
| Bond lengths (Å) | 0.007 | 0.007 |
| Bond angles (°) | 1.3 | 1.4 |
| Ramachandran plot- (%) | |
| Most favored region | 86.0 | 85.0 |
| Additionally favored region | 12.8 | 13.6 |
| Generously favored region | 0.9 | 1.4 |
| Disallowed region | 0.2 | 0.0 |

and Truncate (35). The structure was solved by molecular replacement using the program Phaser (36) with 1QLP (N terminus and RCL removed) as the initial search model. After rigid body refinement with five fragments per monomer, strict NCS was used for several rounds of refinement. Restricted NCS was used in subsequent rounds and later removed entirely. Refinement was conducted using the programs CNS (37) (version 1.0) and Refmac (38). XtalView (39) was used for model building.

**Crystalization, Data Collection, and Refinement for P21**

Crystal Form—Diffraction quality crystals of Δ16 PCI (2 mg/ml in 50% glycerol) were obtained from drops containing 6% PEG3350, 40 mM ammonium citrate, and 10% glycerol. Crystals in 50% glycerol) were obtained from drops containing 6% PEG3350, 40 mM ammonium citrate, and 10% glycerol. Crystals were cryoprotected in 10% PEG3350, 20% glycerol, 40 mM ammonium citrate, and flash-cooled to 100 K. Data were collected at the ESRF (Grenoble, France) station 14-3 from a single crystal and processed as above. The structure was solved by molecular replacement using the refined model from the first crystal form, and refined with CNS. Data collection, processing, and refinement statistics are given in Table 1. Although quality data were obtained to 1.9-Å resolution, completeness was low because of the position of the detector. We therefore truncated the data to 2 Å, although fully complete data were only obtained to 2.2-Å resolution. Figures were created using Pymol, Spock, and Raster3D (40).

**Heparin Binding Studies**—Heparin affinity was initially assessed by elution from a heparin-Sepharose column. Briefly, PCI and its variants were loaded onto a 1 ml of HiTrap heparin Sepharose column (Amersham Biosciences) pre-equilibrated in 50 mM Tris-HCl, 7.4, 50 mM NaCl. The bound protein was then eluted with a 10-column volume gradient from 50 mM to 1 M NaCl, and the conductivity of the peak was measured by the FPLC system (Amersham Biosciences). Accurate dissociation constants were obtained using TNS as an extrinsic fluorescence probe, essentially as described previously for HCII (27). Briefly, precisely size-fractionated heparin (41) was added to buffer (50 mM Tris-HCl, pH7.4, 50 mM NaCl, 0.1% PEG8000, and 1 mM EDTA) containing PCI (200–400 nM) and TNS (10 μM). The change in fluorescence was monitored on a PerkinElmer LS 50B fluorometer with excitation wavelength of 330 nm. Typically, heparin binding resulted in a 6 –10% fluorescence quench at the maximum emission wavelength (~448 nm). When the background TNS fluorescence was subtracted, we observed a 39% quench with a 2–3 nm blue-shift, consistent with previous studies (supplemental Fig. S1) (27, 42). Cleaved PCI had a significantly lower starting fluorescence and a smaller fluorescence change upon titration with heparin (2%). Dissociation constants were determined from the titration data through non-linear regression, as described previously (27). The dependence of binding affinity on heparin length was determined as previously (22) using Equation 1,

$$K_{A_{app}} = K_{A_{int}} \times (N - L + 1)$$  
(Eq. 1)
**Crystal Structure of Native PCI**

| Heparin affinity of the various forms of PCI | Conductivity | $K_d$ with 8mer |
|---------------------------------------------|--------------|----------------|
| PCI                                         | $\mu$S/cm    | $\times 10^{-7}$ n |
| Native plasma                               | 40.8         | $4.6 \pm 0.9$ |
| Native rWT                                  | 53.7         | $1.6 \pm 0.2$ |
| Native P16                                  | 49.6         | $2.7 \pm 0.5$ |
| Native R229A                                | 43.0         | $3.5 \pm 0.4$ |
| Cleaved plasma                              | 43.6         | $5.7 \pm 1.5$ |
| Cleaved rWT                                 | 57.4         | $1.7 \pm 0.7$ |

**Protease Inhibition Assays**—All assays were carried out at room temperature in buffer containing 20 mM Tris-HCl, pH 7.4, 100 mM NaCl, 0.2% bovine serum albumin, 0.1% PEG8000, with 5 mM CaCl$_2$ in the APC assay. For thrombin assays, 5 nM thrombin was incubated with 100 nM PCI and at time intervals, reactions were stopped by adding 300 μM S2238 in the same buffer. For APC assays, 10 nM APC and 100 nM PCI were used, with and without 16.7 μg/ml heparin. Reactions were stopped by adding 600 μM S2366 in the assay buffer, with 0.2 mg/ml polybrene included for assays in the presence of heparin.

**RESULTS**

**Expression and Characterization of Recombinant PCI**—Using plasma-derived native PCI we obtained crystals that diffracted to only 3.4-Å resolution. Further optimization was hindered because the RCL of PCI became cleaved after about 2 weeks, which might have been caused by residual protease activity from the plasma preparation. We thus decided to generate recombinant PCI in *E. coli*. The PCI gene was cloned into pET19b vector and expressed in Rosetta (DE3) pLysS strain, which provides rare codon tRNAs. During purification of recombinant wild-type PCI we discovered a proteolytically susceptible site between His$^{16}$ and Val$^{17}$ (data not shown), so we also expressed an N-terminal-truncated version of PCI starting at Val$^{17}$ (Δ16 PCI). In creating the Δ16 PCI construct, the basic N-terminal region, known as the A+ helix, was removed, allowing us to test the importance of this region for heparin binding and protease recognition. We were eventually able to purify full-length wild-type and the Δ16 PCI variant to homogeneity, and both forms were both SDS-stable complexes with thrombin and APC (data not shown). The rates of thrombin and APC inhibition by wild-type and Δ16 PCI were essentially identical in the absence and presence of catalytic amounts of heparin (data not shown). However, both recombinant forms inhibited thrombin at a rate slightly higher than plasma-derived PCI, as observed previously for *E. coli*-derived PCI (43). Heparin binding affinity, as assessed by elution from heparin-Sepharose, showed a small increase in affinity for the recombinant forms over plasma PCI because of the lack of glycosylation (58), and a marginal decrease in affinity for Δ16 relative to wild-type PCI (Table 2). Based on these initial studies, both forms of recombinant PCI were judged to be active and essentially indistinguishable from native PCI obtained from plasma.

**General Structural Features**—Using native PCI from the Δ16 construct we obtained two different crystal forms. The first form (space group C2) contained three molecules of native PCI per asymmetric unit. This structure was solved to 2.3-Å resolution, and statistics are given in Table 1. Subsequently, a second crystal form (space group P2$_1$) was obtained, leading to a 2.0-Å structure with two copies of native PCI in the asymmetric unit (Table 1). Monomers from the first crystal form are denoted C2_A, B, and C, and those from the second form are denoted P2$_1$/A and B. The five structures were quite similar (average C$_\alpha$ R.m.s.d. $\sim$1 Å), and superposition of the monomers provided an NMR-like ensemble of structures to reveal the regions of inherent flexibility (Fig. 1A). Flexibility is notable in the RCL, the helix D region and the very N terminus. The localized flexibility apparent from this superposition is supported by higher average temperature factors in these regions. The overall fold is that of a typical serpin, with a superposition of α$_1$-antitrypsin resulting in a C$_\alpha$ R.m.s.d. of $\sim$1.5 Å for 321 equivalent atoms; however, some significant differences are observed (Fig. 1B). Of particular interest is the relative accessibility of the RCL and the reactive center P1 Arg residue, the rotation of helices D and H, and the two turn shortening of helix A. A similar superposition with TBG (44) is shown in Fig. 1C, and although the sequence identity is lower, the structures are more similar, with the positions of helices D and H preserved. The functional consequence of this structural feature is discussed below.

The RCL—Based on structural alignment with the prototypical serpin α$_1$-antitrypsin, the RCL of PCI is 3–4 residues longer (depending on the assignment of the first residue in s1C) through an insertion C-terminal to the P1 residue. While this is also true for some other serpins such as AT, in PCI the extra residues confer a degree of flexibility never before seen in a native serpin. Of the five molecules of native PCI in our two crystal forms, the RCL is only fully modeled in one (P2$_1$/A), however, the density is poor from P6-P9. For all other monomers the RCL was only partially modeled into electron density (18 missing from P2$_1$/B, 7 from C2_A, 4 from C2_B and 17 from C2_C), and the ability to model RCL residues was dependent on the presence of crystal contacts. The reactive center residue Arg$^{357}$ was only observed in one of the five structures, indicating the high degree of flexibility of this region. The portion of the RCL which is recognized by proteases (generally P4–P3$^\prime$) is not in a fixed canonical β-sheet like conformation as seen for other serpins (45, 46). The RCL of PCI is thus an unusually long and flexible loop, which serves as a pseudosubstrate for its targets, and unlike other heparin-binding serpins, AT and HCII, release of the hinge region from β-sheet A is not part of the heparin activation mechanism.

The N-terminal portion of the RCL (the hinge region) is seen in all five crystal structures of native PCI (Fig. 2). The conformation of the hinge is identical in four structures, with main chain β-sheet contacts between residues P16–P13 and the loop following s3A (190–193), and this would appear to be the preferred conformation of the native form. β-Sheet A is zipped up to the normal extent for a native serpin, with the final β-sheet H-bond between the O of residue 186 (s3A) and the N of residue 335 (s5A) (Fig. 2, A and B). As is usually observed in native serpins, there is a small gap between strands 3 and 5 at the top of β-sheet A. This gap is filled in four structures by two conserved water molecules (Fig. 2B). Interestingly, the conformation of the side chain of Lys$^{187}$ is also conserved in these four structures.
of native PCI. Lys187 is located on s3A and its basic head-group interposes between two acidic residues on strand 5A (Glu335 and Asp337). These interactions cross-link strands 3 and 5 of 8-sheet A, and, together with the hinge region interactions (not shown in the figure), serve to limit premature loop incorporation which might otherwise be expected for such a long and flexible RCL.

However, for one of the five structures of native PCI the conformation of the hinge and shutter regions suggest an equilibrium involving partial hinge region incorporation into 8-sheet A. In the third monomer of the C2 crystal form (C2_C) we observed an opening at the top of sheet A similar to the extent seen in the crystal structure of murine antichymotrypsin (47) (Fig. 2, C and D). The last 8-sheet hydrogen bond between s3A and s5A forms two residues further down the sheet (O from 184 and N from 333) than seen in the four other native structures. In murine antichymotrypsin the P15 Gly residue is participating in 8-sheet hydrogen bonds and can thus be considered to be incorporated into sheet A. Although sheet A is opened to the same extent in this form of PCI, the P15 Gly is not inserted. Rather, the hinge region has displaced the side chain of Lys187 and appears to be poised for insertion. The movement of Lys187 breaks the bridge between strands 3 and 5A, and also engenders a repulsive contact between the side chain of Glu335 and the O of 186 which would predictably aid in the separation of the strands to allow RCL incorporation. Thus, this conformation of PCI may reveal the events which occur just prior to hinge region incorporation and consequent RCL insertion into 8-sheet A. A video depiction of this important serpin conformational event is given as supplementary material (see supplemental video).

**Heparin-binding Site and Properties**—Based on the structure of cleaved PCI we proposed an extended heparin-binding site encompassing an extensive basic patch including helix H and residues on sheet C (Arg229, Arg234, and Arg362) (34). The conformation of this region in the structures of native PCI is indistinguishable from that of the cleaved form, consistent with the observation that native and RCL-cleaved PCI bind equally well to heparin-Sepharose (18) (Table 2). Traditionally, heparin affinity is assessed by elution from a heparin-Sepharose column when subjected to a salt gradient. We compared the apparent affinities of plasma and recombinant PCI in such a fashion, and the elution conductivities are given in Table 2.

To obtain accurate dissociation constants for PCI and its variants, we adapted a method used for other heparin-binding serpins whereby a change in quantum yield of the extrinsic orange). PCI is shown in the classic orientation with the RCL on top and 8-sheet A facing. The P1 residue could only be modeled in P21_A (yellow ball) as the RCL is highly flexible in the other structures. B, ribbon depiction of PCI (gray with 8-sheet A red, the RCL yellow and orange, helix A green, and helix H blue) is shown superimposed with the prototypical serpin ut-Antitrypsin (semitransparent cyan). The two molecules are rotated ~120° to the left on the long axis to highlight important structural differences. The RCL is in a more exposed and flexible conformation in PCI, due in part to the 3–4 residue insertion (orange) C-terminal to the P1 residue (rods). The two-turn shortening of helix A is also evident, as it the rotation of helix H. C, result of these conformational changes is the exposure of a hydrophobic pocket similar to that observed for TBG (semitransparent magenta). Thyroxine (magenta rods) binds to a hydrophobic pocket formed in TBG by a similar rotation of helix H (colored as before).
fluorescence probe TNS is followed during titration with heparin. As for AT (42) and HCII (27), heparin binding to PCI resulted in a fluorescence quench (supplemental Fig. S1). Thus we were able to obtain accurate dissociation constants for the various forms of PCI, which supported the data obtained using heparin-Sepharose (Table 2). We also applied this method to determine the dependence of affinity on heparin chain length (Fig. 3 and Table 3). The minimum heparin length capable of fully occupying the binding site on PCI was calculated from the linear fit of association constants versus heparin length (Fig. 3B). The calculated value of 8 monosaccharide units per the minimum site size corresponds well to the observed loss of affinity for the 6mer fragment (Fig. 3A), and the slope yields an intrinsic dissociation constant of 190 nM under these conditions.

Using this method we were able to demonstrate that the N-terminal tail does not appreciably affect the heparin affinity of PCI (1.6 versus 2.7 \times 10^{-7} M, Table 2). To test if heparin binds exclusively to helix H or across the whole basic patch, we generated the PCI variant R229A (Fig. 4A) and determined its affinity for the heparin 8mer. As heparin affinity for the variant was only marginally reduced (2.7 versus 3.5 \times 10^{-7} M), we concluded that Arg229 is not part of the heparin-binding site of PCI. These data combined with previous mutagenesis studies suggest that heparin binding occurs exclusively along helix H, and that the apparent involvement of other regions reflects nonspecific electrostatic effects. Thus, we propose a mode for heparin binding in which the short heparin 8mer binds along the length of helix H (Fig. 4B).

**Hydrophobic Ligand-binding Site**—One of the surprise features of our previous structure of RCL-cleaved PCI was the two-turn N-terminal shortening of helix A. As mentioned above, this feature is fully preserved in all of the native structures presented here (Fig. 1). The shortening of helix A, coupled with the rotation of helix H created a large hydrophobic pocket we termed the helix A gap, and this feature is also present in the...
Crystal Structure of Native PCI

PCI is a multifunctional serpin with wide ranging protease inhibitory activity. How specificity can be conferred to the highly promiscuous serpin mechanism is one of the outstanding questions in the field. AT and HCII achieve specificity by constraining their RCLs to limit proteolytic accessibility and by hiding critical exosites until heparin binds (25, 49, 50). It was conceivable that PCI, as a heparin-binding serpin, used an analogous mechanism. However, the structures of native PCI described here reveal a fully exposed and flexible RCL. Indeed, the RCL of PCI is longer than that of a typical serpin by 3–4 residues inserted C-terminal to the P1 Arg, resulting in an unusual degree of accessibility of the region generally recognized by target proteases (P4–P3'). PCI thus differs from its related heparin-binding serpins by circulating in a conformation permissive to proteolytic attack. GAG binding is unlikely to induce a conformational change in PCI, and may only serve to localize PCI to surfaces where it can encounter target proteases.

How then can specificity be conferred? Is it only a matter of co-localization? One thing which has become clear with the solution of crystal structures of serpins with target proteases is that exosites play a critical role in determining specificity. In fact, all of the structures of physiologically relevant Michaelis complexes show exosite interactions between the serpin and the protease. These are: HCII and thrombin (25), AT and thrombin (51), and AT and factor Xa (50). Interestingly, the RCL played a key role in the availability of exosites on the surface of the two proteins, because it must be engaged in the active site of the protease. Exosite I of thrombin docks toward the back of HCII in the Michaelis complex, and this interaction is only possible with the extension of the hinge region out of, and away from β-sheet A. A similar dependence on hinge region extension for the recognition of factors IXa and Xa was demonstrated by biochemical methods (52). Soon after, the structure of the Michaelis complex between AT and factor Xa revealed how the orientation of factor Xa on AT necessitated the extension of the hinge region. In contrast, docking of thrombin on AT is insensitive to extension of the hinge region; rather, exosite contacts can only be formed through extension of the C-terminal portion of the RCL. AT, like PCI, has a three residue RCL insertion C-terminal to the reactive center residue, however, in AT this insertion forms a tight hydrogen-bonded turn and does not increase the flexibility of the RCL. In the crystal structure of the AT–thrombin Michaelis complex the hydrogen bonds are broken and the P region is stretched to allow exosite contacts toward the front of AT (51).

These complexes are relevant to PCI in that they demonstrate the role that RCL flexibility/extensibility on the N- and C-terminal sides can play in determining permissible exosite contact surfaces, and thus in determining target protease specificity. It should be noted that PCI in the absence of cofactor is actually a poor inhibitor of its known targets, and this may well be caused by the high degree of flexibility of its RCL. Most native serpin structures show a canonical RCL conformation in

TABLE 3  
The effect of heparin size on binding affinity

| Heparin size | Dissociation constant (M⁻¹) |
|--------------|----------------------------|
| 6            | 8.0 ± 1.0 × 10⁻⁶           |
| 8            | 2.7 ± 0.5 × 10⁻⁷           |
| 10           | 1.5 ± 0.1 × 10⁻⁷           |
| 12           | 4.0 ± 0.7 × 10⁻⁸           |
| 14           | 3.0 ± 0.6 × 10⁻⁸           |

in TBG, lending support to the contention that this pocket is the hydrophobic ligand-binding site of PCI.

DISCUSSION

native structure of PCI (Fig. 4C). Based on molecular docking studies we proposed that this hydrophobic pocket is the retinoic acid-binding site of PCI. This hypothesis was strengthened by the recently solved structure of TBG bound to its hydrophobic ligand thyroxine (44) (Figs. 1C and 4, C and D). Thyroxine is much shorter than retinoic acid and consequently its binding site is formed by the rotation of helix H alone. When TBG is superimposed on native PCI, thyroxine fits comfortably into the top part of the helix A gap (Fig. 4, C and D). In the P2₁ crystal of native PCI, one of the two molecules (B) has its helix A gap filled by the hydrophobic portion of the RCL of the other monomer (A). As shown in Fig. 4, C and D, the position of RCL residues from P5 to P1 (IFTFR) corresponds to that of thyroxine...
and around the P1 residue (46). The 3–4 residue insertion C-terminal to the reactive center lifts the P1 away from the body of PCI and induces flexibility. There is thus an entropic cost for docking the RCL into the active site of the protease, predictably resulting in an increased dependence on exosite interactions for protease recognition.

Another unusual feature of PCI is the use of helix H instead of helix D as the principal heparin-binding site. PCI is a highly basic molecule (PI for non-glycosylated form, 9.4 with net charge of 8) and prone to nonspecific interactions with heparin. Thus the true extent of the binding site has been difficult to determine. It is clear from the surface charge distribution shown in Fig. 4A that helix H is contiguous with a large basic patch composed of residues from β-sheet C. In addition, wild-type PCI possesses a basic N-terminal tail which has been implicated in heparin binding. To determine the minimal heparin binding size and to allow for a more accurate identification of heparin binding regions, we developed a sensitive method for determining the affinity of PCI for heparin fragments. We demonstrated that the minimal heparin binding size for PCI is 8 monosaccharides in length. This size roughly corresponds to the diameter of the basic region near helix H. To determine if residues in the extended basic patch are utilized in heparin binding we mutated a solvent exposed residue, Arg229, to Ala. This mutation was found to have little effect on affinity for the 8mer heparin fragment, despite a reduction in the ionic strength required for elution from heparin-Sepharose. Similarly, we showed that the N-terminal 16 amino acids, including the so-called A+ helix, do not contribute significantly to heparin binding (Table 2).

Our data also support the previous observation that heparin affinity is independent of cleavage in the RCL (18). The modeled position of the heparin 8mer bound to PCI is of clear functional relevance, as it places heparin where it is able to directly contact proteases in Michaelis complexes. The length and flexibility of the RCL of PCI, particularly the C-terminal 3–4 residue insertion, allows the protease to dock toward helix H, and such an orientation may juxtapose the heparin-binding sites of the protease and PCI. In this case, heparin would not be bridging the two proteins in a traditional fashion, but would effectively create a new exosite on the surface of PCI. Superposition of PCI and APC on deposited serpin-protease Michaelis complex structures provides one such orientation. A model of the PCI-heparin-APC Michaelis complex based on the structure of AT with factor Xa approximates the known heparin-binding site (30 and 60 loops of APC (53–55)) with the modeled position of heparin on PCI. While this is not likely to provide any useful details relating to individual contacts, it nicely illustrates the potential for such a heparin sandwiching mechanism as proposed previously (34). However, further study is required.

FIGURE 4. Surface properties of PCI reveal the heparin and hydrophobic ligand-binding sites. A, electrostatics of native PCI, oriented as in Fig. 1B, are depicted as a semitransparent surface (blue is positive and red is negative). The underlying residues on helix H are shown, as is the adjacent Arg229, revealing the extent of the possible heparin-binding site of PCI. B, because mutation of Arg229 did not affect the affinity of PCI for size-fractionated heparin, we conclude that only residues along the face of helix H are involved in the heparin interaction. We thus can model the minimal 8mer heparin fragment on helix H, outside contact distance for Arg229 (close up of A). C, hydrophobic properties of native PCI are similarly visualized, with hydrophobic patches colored green, with the hydrophobic region of the RCL indicated by the oval. The position of thyroxine is shown as magenta rods, and the observed crystallographic contacts involving the RCL are in yellow. D is a close-up of the hydrophobic pocket and illustrates the conservation of hydrophobic binding sites between TBG and PCI.
to determine the precise nature of the interaction between PCI and heparin, and the molecular basis of improved protease recognition.

In addition to the broad inhibitory specificity of PCI and its regulation by cofactor binding, PCI is of interest for its ability to bind to hydrophobic ligands such as retinoic acid. This ability is shared by the non-inhibitory hormone transporting serpins CBG and TBG. Although PCI is unlikely to transport hormones in the circulation, such a function may be relevant in other tissues. Our earlier structure of RCL-cleaved PCI gave some clues as to the location of the hydrophobic hormone-binding site, which we termed the helix A gap. As PCI was the first hormone-binding serpin with a resolved crystal structure, it was natural to entertain the possibility that this gap would also be found in CBG and TBG. Recently, the structure of TBG bound to thyroxine was solved by Zhou et al. (44), and the observed binding site corresponded nicely with our original hypothesis. While in PCI the hydrophobic site is partially created by the two-turn shortening of helix A, the conserved rotation of helix H and small shift in helix D are sufficient to create the thyroxine-binding site of TBG (Fig. 1C). The larger hydrophobic pocket found in PCI may be necessary to accommodate the longer and more hydrophobic retinoids, or might be utilized for the apparently specific interaction with the phospho-pocket found in PCI may be necessary to accommodate the thyroxine-binding site of TBG (Fig. 1C). The larger hydrophobic pocket found in PCI may be necessary to accommodate the longer and more hydrophobic retinoids, or might be utilized for the apparently specific interaction with the phospho-

REFERENCES
1. Geiger, M., Zechmeister-Machhart, M., Uhrin, P., Hufnagl, P., Ecke, S., Priglinger, U., Xu, J., Zheng, X., and Binder, B. R. (1996) Immunopharmacology 32, 53–56
2. Gettins, P. G. W., Patston, P. A., and Olson, S. T. (1996) Serpins: Structure, Function and Biology. R. G. Landes Co., Austin, TX
3. Gettins, P. G. (2002) Chem. Rev. 102, 4751–4804
4. Marlar, R. A., and Griffin, J. H. (1980) J. Clin. Invest. 66, 1186–1189
5. Suzuki, K., Nishioaka, J., and Hashimoto, S. (1983) J. Biol. Chem. 258, 163–168
6. Rezaie, A. R., Cooper, S. T., Church, F. C., and Esmon, C. T. (1995) J. Biol. Chem. 270, 25336–25339
7. Elisen, M. G., von dem, B., Bouma, B. N., and Meijers, J. C. (1998) Blood 91, 1542–1547
8. Geiger, M., Krebs, M., Jerabek, L., and Binder, B. R. (1997) Immunopharmacology 36, 279–284
9. Huntington, J. A., Read, R. J., and Carrell, R. W. (2000) Nature 407, 923–926
10. Olson, S. T., Swanson, R., Day, D., Verhamme, I., Kvassman, J., and Shore, J. D. (2001) Biochemistry 40, 11742–11756
11. Huntington, J. A. (2006) Trends Biochem. Sci. 31, 450–455
12. Schechter, I., and Berger, A. (1967) Biochem. Biophys. Res. Commun. 27, 157–162
13. Prendes-Garcia, M. J., Zechmeister-Machhart, M., Bielek, E., Vanyek-Za-
vadil, E., Breuss, J., Binder, B. R., and Geiger, M. (2001) J. Thromb. Haem.
July, 0X 989 (Abstract)
14. Dementie, A., Dobbo, J., and Gettins, P. G. (2006) J. Biol. Chem. 281, 3452–3457
15. Komiyama, T., Gron, H., Pemberton, P. A., and Salvesen, G. S. (1996) Protein Sci. 5, 874–882
16. Komiyama, T., Ray, C. A., Pickup, D. J., Howard, A. D., Thornberry, N. A., Peterson, E. P., and Salvesen, G. (1994) J. Biol. Chem. 269, 19331–19337
17. Irving, J. A., Pike, R. N., Dai, W., Bromme, D., Worrall, D. M., Silverman, G. A., Coetzter, T. H., Dennison, C., Bottomley, S. P., and Whistock, J. C. (2001) Biochemistry 41, 4985–4991
18. Pratt, C. W., and Church, F. C. (1992) J. Biol. Chem. 267, 8789–8794
19. Garg, H. G., Linhardt, R. J., and Hales, C. A. (2005) in Chemistry and Biology of Heparin and Heparan Sulfate (Garg, H. G., Linhardt, R. J., and Hales, C. A., eds) pp. 367–398, Elsevier, Oxford
20. Kuhn, L. A., Griffin, J. H., Fisher, C. L., Greengard, J. S., Bouma, B. N., Espana, F., and Tainer, J. A. (1999) Proc. Natl. Acad. Sci. U. S. A. 87, 8506–8510
21. Shirk, R. A., Elisen, M. G., Meijers, J. C., and Church, F. C. (1994) J. Biol. Chem. 269, 28690–28695
22. Neese, L. L., Wolfe, C. A., and Church, F. C. (1998) Arch. Biochem. Bio-
phys. 355, 101–108
23. Huntington, J. A. (2003) J. Thromb. Haemost. 1, 1535–1549
24. Carrell, R. W., Stein, P. E., Fermi, G., and Wardell, M. R. (1994) Structure 2, 257–270
25. Baglin, T. P., Carrell, R. W., Church, F. C., Esmon, C. T., and Huntington, J. A. (2002) Proc. Natl. Acad. Sci. U. S. A. 99, 11079–11084
26. Jin, L., Abrahams, J. P., Skinner, R., Pettit, M., Pike, R. N., and Carrell, R. W. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 14683–14688
27. O’Keeffe, D., Olson, S. T., Gasiunas, N., Gallagher, J., Baglin, T. P., and Huntington, J. A. (2004) J. Biol. Chem. 279, 50267–50273
28. Shen, L., Villoutreix, B. O., and Dahlback, B. (1999) Thromb. Haemost. 82, 72–79
29. Ecke, S., Seiger, M., Resch, I., Jerabek, I., Sting, L., Maier, M., and Binder, B. R. (1992) J. Biol. Chem. 267, 7048–7052
30. Hammond, G. L., Smith, C. L., Goping, I. S., Underhill, D. A., Harley, M. J., Reventos, J., Musto, N. A., Gunsalus, G. L., and Bardin, C. W. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 5153–5157
31. Pemberton, P. A., Stein, P. E., Pepys, M. B., Potter, J. M., and Carrell, R. W. (1988) Nature 336, 257–258
32. Suda, S. A., Gettins, P. G., and Patston, P. A. (2000) Arch. Biochem. Bio-
phys. 384, 31–36
33. Jerabek, I., Zechmeister-Machhart, M., Binder, B. R., and Geiger, M. (2001) Eur. J. Biochem. 268, 5989–5996
34. Huntington, J. A., Kjellberg, M., and Stenflo, J. (2003) Structure 11, 205–215
35. Leslie, A. W. G. (1992) Joint CCP4 and ESF-EACMB Newsletter on Protein Crystallography. Daresbury Laboratory, Warrington, UK, Vol. 26
36. McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C., and Read, R. J. (2005) Acta Crystallography. D Biol. Crystallography 61, 458–464
37. Brunger, A. T., Adams, P. D., Clore, G. M., Trewhella, T., Wuthrich, K., and Bock, A. J. (1998) Acta Crystallography C 54, 905–921
38. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Acta Crystallogr. Sect. D 53, 240–255
39. McRee, D. E. (1992) J. Mol. Graphics 10, 44–46
40. Merritt, E. A., and Murphy, M. E. P. (1994) Acta Crystallogr. Sect. D 50, 869–873
41. Goger, B., Halden, Y., Rek, A., Mosl, R., Pye, D., Gallagher, J., and Kungl, A. J. (2002) Biochemistry 41, 1640–1646
42. Meagher, J. L., Olson, S. T., and Gettins, P. G. (2000) J. Biol. Chem. 275,
Crystal Structure of Native PCI

43. Rehault, S. M., Zechmeister-Machhart, M., Fortenberry, Y. M., Malleier, J., Binz, N. M., Cooper, S. T., Geiger, M., and Church, F. C. (2005) Biochim. Biophys. Acta 1748, 57–65
44. Zhou, A., Wei, Z., Read, R. I., and Carrell, R. W. (2006) Proc. Natl. Acad. Sci. U. S. A. 103, 13321–13326
45. Elliott, P. R., Abrahams, J. P., and Lomas, D. A. (1998) J. Mol. Biol. 275, 419–425
46. Dementiev, A., Simonovic, M., Volz, K., and Gettins, P. G. (2003) J. Biol. Chem. 278, 37881–37887
47. Horvath, A. J., Irving, J. A., Rossjohn, J., Law, R. H., Bottomley, S. P., Quinsey, N. S., Pike, R. N., Coughlin, P. B., and Whisstock, J. C. (2005) J. Biol. Chem. 280, 43168–43178
48. Elisen, M. G., Maseland, M. H., Church, F. C., Bouna, B. N., and Meijers, J. C. (1996) Thromb. Haemost. 75, 760–766
49. Johnson, D. J., Langdown, J., Li, W., Luis, S. A., Baglin, T. P., and Huntington, J. A. (2006) J. Biol. Chem. 281, 35478–35486
50. Johnson, D. J., Li, W., Adams, T. E., and Huntington, J. A. (2006) EMBO J. 25, 2029–2037
51. Li, W., Johnson, D. J., Esmon, C. T., and Huntington, J. A. (2004) Nat. Struct. Mol. Biol. 11, 857–862
52. Langdown, J., Carter, W. J., Baglin, T. P., and Huntington, J. A. (2006) FEBS Lett. 580, 4709–4712
53. Glasscock, L. N., Gerlitz, B., Cooper, S. T., Grinnell, B. W., and Church, F. C. (2003) Biochim. Biophys. Acta 1649, 106–117
54. Friedrich, U., Blom, A. M., Dahlback, B., and Villoutreix, B. O. (2001) J. Biol. Chem. 276, 24122–24128
55. Yang, L., Manithody, C., and Rezaie, A. R. (2002) Biochemistry 41, 6149–6157
56. Nishioka, J., Ning, M., Hayashi, T., and Suzuki, K. (1998) J. Biol. Chem. 273, 11281–11287
57. Fulton, K. F., Buckle, A. M., Cabrita, L. D., Irving, J. A., Butcher, R. E., Smith, I., Reeve, S., Lesk, A. M., Bottomley, S. P., Rossjohn, J., and Whisstock, J. C. (2005) J. Biol. Chem. 280, 8435–8442
58. Fujita, M., Izutani, W., Takahashi, K., Nishizawa, K., Shirono, H., and Koga, I. (2002) Thromb. Res. 105, 95–102