The High Resolution Crystal Structure of the Human Tumor Suppressor Maspin Reveals a Novel Conformational Switch in the G-helix*

Received for publication, October 25, 2004, and in revised form, March 8, 2005 Published, JBC Papers in Press, March 10, 2005, DOI 10.1074/jbc.M412043200

Maspin is a serpin that acts as a tumor suppressor in a range of human cancers, including tumors of the breast and lung. Maspin is crucial for development, because homozygous loss of the gene is lethal; however, the precise physiological role of the molecule is unclear. To gain insight into the function of human maspin, we have determined its crystal structure in two similar, but non-isomorphous crystal forms, to 2.1- and 2.8-Å resolution, respectively. The structure reveals that maspin adopts the native serpin fold in which the reactive center loop is expelled fully from the A β-sheet, makes minimal contacts with the core of the molecule, and exhibits a high degree of flexibility. A buried salt bridge unique to maspin orthologues causes an unusual bulge in the region around the D and E α-helices, an area of the molecule demonstrated in other serpins to be important for cofactor recognition. Strikingly, the structural data reveal that maspin is able to undergo conformational change in and around the G α-helix, switching between an open and a closed form. This change dictates the electrostatic character of a putative cofactor binding surface and highlights this region as a likely determinant of maspin function. The high resolution crystal structure of maspin provides a detailed molecular framework to elucidate the mechanism of function of this important tumor suppressor.

Maspin (mammary serine proteinase inhibitor (SERPINB5)) was initially identified as a tumor-suppressing serpin down-regulated in invasive mammary carcinoma cell lines (1). Maspin loss in numerous cancers (including breast, prostate, squamous cell carcinoma, gastric cancer, and lung) correlates with metastasis and a poor clinical prognosis (for a review, see Ref. 2). In contrast, high levels of maspin expression in certain cancers (in particular, pancreatic and ovarian cancer) correlate with tumor invasion and poor survival. Like other clade B serpins (3), maspin has a nucleocytoplasmic distribution, however it is also found at the cell surface (1, 4, 5). The intracellular role of maspin is at present unclear, but it has been suggested to play a role in apoptosis pathways (6). A large body of evidence suggests that maspin has an important extracellular role: it can suppress tumor growth and metastasis in vivo and tumor cell motility and invasion in vitro (1, 7, 8). Maspin also plays a fundamental role in early embryonic development; murine knock-out studies reveal that it is essential for proper organization of the epiblast (9). Consistent with a complex role in tumorigenesis, maspin also exhibits anti-angiogenic activity (10), and expression of the maspin gene has been demonstrated to be under the control of the oncogenic transcription factors p53 and p63 (11, 12). Because the failure to properly control proteolytic activity can result in disruption of the basement membrane and promote tumor invasion, it was initially hypothesized that maspin may exert its anti-metastatic effect by functioning as a pericellular protease inhibitor (1).

Serpins are unusual molecules that fold into a metastable native state. The majority of serpins function to inhibit either serine or cysteine proteases; however, non-inhibitory members of the family have also been identified, including the avian serpin ovalbumin and the human angiogenesis inhibitor pigment epithelium-derived factor (PEDF) (3, 13, 14). All known inhibitory serpins inactivate target proteases via a major conformational change termed the stressed (S) to relaxed (R) transition. During this rearrangement, the region responsible for docking with the protease (the reactive center loop (RCL)) is cleaved and inserts into the center of a large β-sheet (the A β-sheet), forming an extra β-strand. This conformational rear-

* This work was supported in part by the National Health and Medical Research Council of Australia (NHMRC), the Australian Research Council, and the state government of Victoria (Australia). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (codes 1XU8 and 1WZ9) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

† These authors contributed equally to this work.

‡ An Anti Cancer of Victoria Research Fellow and Monash University Research Fellow.

§ A Monash University Logan Fellow and R. D. Wright Fellow of the NHMRC.

∥ Supported by a Wellcome Trust Senior Research Fellowship in Biomedical Science in Australia.

¶ Joint senior authors.

To whom correspondence may be addressed: Tel.: 613-9405-3736; Fax: 613-9905-4699; E-mail: Jamie.Rossjohn@med.monash.edu.au.

An NHMRC Senior Research Fellow and Monash University Senior Logan Fellow. To whom correspondence may be addressed: Tel.: 613-9905-3747; Fax: 613-9905-4699; E-mail: James.Whisstock@med.monash.edu.au.

1 The abbreviations used are: PEDF, pigment epithelium-derived factor; r.m.s.d., root mean square deviation; RCL, reactive center loop; BisTris, 2-[bis(2-hydroxyethyl)aminol]-2-(hydroxyethyl)propane-1,3-diol; PAI-2, plasminogen activator inhibitor-2.
rangement is responsible for the translocation and trapping of the target protease (15). Mutations within the RCL that interfere with the S to R transition (for example, within the conserved hinge region) abolish or seriously compromise inhibitory activity. Consistent with these data, serpins (such as ovalbumin and PEDF) that do not function as protease inhibitors have non-conserved hinge regions and lack the ability to undergo the S to R transition under physiological conditions (16).

Biochemical and biophysical studies reveal that maspin does not undergo the S to R transition and is thus unable to inactivate proteases in a classic serpin-like fashion (17). Nor does the RCL of maspin contain the consensus hinge region motif present in inhibitory serpins (18, 19). It has been suggested that maspin may control the urokinase-type plasminogen activator and/or tissue-type plasminogen activator; however, recent studies do not support this view (8, 20). Interestingly, however, the RCL of maspin is clearly important for function: studies using synthetic maspin RCL peptides as well as maspin/ovalbumin chimeras reveal that this region is important for promoting cell adhesion (8, 21). Taken together, these data suggest that maspin is a non-inhibitory serpin that functions to suppress angiogenesis and metastasis through an as yet uncharacterized protein-protein or protein-ligand interaction.

The role of maspin in tumor progression, angiogenesis, and embryogenesis is of great interest, but many aspects of maspin biology, such as the underlying biochemical reasons behind the apparently conflicting role of maspin in different cancers, remain to be understood. To begin to understand maspin at the molecular level we have determined two crystal structures of this important human tumor suppressor, at 2.1- and 2.8-Å resolution. Recently a 3.1-Å crystal structure of a maspin mutant lacking cysteine residues has been reported (22); however, the low resolution of this structure places a limit on structural interpretation. High resolution structures of the wild-type protein reveal that maspin is capable of undergoing conformational change in and around the G-helix, a region of the serpin scaffold previously understood to be conformationally inert. In particular, structural changes in this region result in a major reorganization of charged residues and the manifestation of a large negatively charged patch centered on the G-helix. We suggest that this region represents a cofactor binding site under conformational control.

EXPERIMENTAL PROCEDURES

Protein Expression and Purification—The pRSET/maspin plasmid (21) was modified by restriction enzyme digest, using NdeI and EcoRI to excise the translated codons upstream of the maspin cDNA and replace them with the following oligonucleotides after annealing: 5′-tataggagggttctctcatcatcatcatcatcatgaaaacctgtattttcagggccag and 5′-aattctggccctcctcatcatcatcatcatcatgaaaacctgtattttcagggccag and this material was used throughout the course of the work. Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (Voyager-DE STR Biospectrometry Workstation, Applied Biosystems, Framingham, MA) revealed that the molecular mass of the purified product was 42.92 kDa, compared with a theoretical molecular mass of 42.53 kDa. Because β-mercaptoethanol was used in all buffers throughout the purification and crystallization steps, the molecular weight discrepancy corresponds to approximately five cysteine residues derivatized with β-mercaptoethanol. The derivation of cysteine residues is consistent with the high resolution crystal structures, which show three and six derivatized cysteines for each of the two molecules in the asymmetric unit.

Maspin Cleavage by Cathepsin L—Recombinant maspin (18 μM) was treated with 0.25 μM recombinant human cathepsin L (24) at 37 °C for 30 min in 0.1 M NaAc, pH 5.5, 1 mM EDTA, 0.1% Brij-35, 10 mM cysteine. Complete cleavage within the RCL was confirmed by SDS-PAGE, and the molecular masses of the cleaved products were found to be 38.24 and 4.707 kDa by mass spectrometry. Based on these data (and the observation that the only cysteine residue in the C-terminal peptide is buried and unmodified in all crystal structures) the position of the cleavage site was mapped to the maspin RCL.

Spectroscopic Methods—Circular dichroism experiments were performed using a Jasco 810 Spectropolarimeter (Jasco, Tokyo). The protein concentration for both cleaved and native maspin was used 15 μM with a 0.1-cm path length. Thermal denaturation was performed at a heating rate of 1 °C/min in 20 mM Tris-HCl, pH 8.0, and monitored at 222 nm.

Cell-Extracellular Matrix Adhesion Assay—The breast carcinoma cell line MDA-MB-231 (provided by T. Brown, Monash University) was routinely cultured in RPMI 1640 medium containing 10% fetal calf serum and 2 mM glutamine (Invitrogen). For the cell adhesion assay, wells were coated with 50 μl of 5 μg/ml fibronectin (Sigma) and blocked with 2% bovine serum albumin in phosphate-buffered saline. The cell adhesion assay was performed essentially as described by Ngaamkit-suchakul et al. (8), except that prior to assessing adhesion, cells in suspension were reincubated with the indicated concentration of recombinant maspin for 30 min at 37 °C, then directly added to the wells.

Crystallization—Maspin was concentrated to 20 mg/ml and a Carlsberg Honeybee crystallization robot (Genomic Solutions) was used to establish initial crystallization conditions (100-nl drops). Small needles were identified in a condition containing 0.1 M Tris, pH 8.28, 3 M (NH4)2SO4. Subsequent fine screening was carried out using the hanging drop vapor diffusion method. Single large crystals grew at 22 °C in 0.1 M BisTris, pH 8.3, 2.9 M (NH4)2SO4 after 4–6 weeks. The crystals were flash-frozen prior to data collection with 5% ethylene glycol as the cryoprotectant.

X-ray Data Collection, Structure Determination, and Refinement—The first crystal form of maspin diffracted to 2.8-Å resolution and belongs to space group P212121, with unit cell dimensions of a = 53.54 Å, b = 100.48 Å, c = 139.23 Å. The second crystal form diffracted to 3.5-Å resolution and belongs to space group P2221, with unit cell dimensions of a = 54.12 Å, b = 95.05 Å, c = 139.23 Å. Both structures are consistent with two monomers in the asymmetric unit and ~42% solvent content (calculated using Matthews (25)). The data were integrated and scaled with the HKL suite (26). See Table I for a summary of data collection statistics.

The 2.8-Å structure was solved using an implementation of the molec- ular replacement method (resolution range, 10–3.5 Å) within the PHASE program (27) and multiple superposed search models of native serpins (ovalbumin, 10VA; antitrypsin, 1QLP; serpin 1K, 1SEK, antithrombin, 1E05). Two clear peaks in the rotation function yielded two solutions in the translation function that packed well within the unit cell and, together with unbiased features in the initial electron density map, confirmed the correctness of the monoclinic replacement solution.

An initial model was built using native ovalbumin (10VA), with all sequence differences mutated to alanine. The progress of refinement was monitored by the Rfree value (5% of the data) with neither a sigma, nor a low resolution cut-off applied to the data. The structure was refined using rigid-body fitting of the individual domains followed by the simulated-annealing protocol implemented in CNS (version 1.1) (28), iterated with rounds of model building using the program O’ (29). Tightly restrained individual B-factor refinement was employed, and bulk solvent correction was applied to the data set, which led to a significant drop in the Rfree value. Further refinement, incorporating translation, libration, and screw-rotation displacement refinement was then carried out using REFMAC (30). A bulk solvent correction (Babinet model with mask) was also used within REFMAC. Throughout refinement, tight non-crystallographic symmetry restraints were imposed on the two molecules, excluding residues involved in crystal contacts. Water molecules were included in the model if they were within hydrogen-bonding distance to chemically reasonable groups, if they appeared in Fc − Fo maps contoured at 3.5σ and had a B-factor <60 Å2. The RCL loop in both monomers was observed to be mobile and refined as a disordered and modeled fully as a non-crystallographic symmetry restraint.

On obtaining better diffracting crystals, the 2.1-Å structure was solved by molecular replacement using AMORE (31) and chain A of the refined 2.8-Å model as a search probe. Subsequent structure refinement proceeded as for the 2.8-Å structure using CNS, however the higher resolution data permitted full positional and B-factor refinement without non-crystallographic symmetry restraints. Further refinement, incorporating translation, libration, and screw-rotation displacement re-
Fig. 5 (A) and (B) going from blue (potential of 28.3 kcal/mol) to red (potential of 7.6 kcal/mol). Other residues were considered neutral. The calculation was done as described in the Materials and Methods section. Lys and Arg residues were assigned a single positive charge, while other charged residues were considered neutral. The Poisson-Boltzmann solver was used to calculate the static potential. GRASP (35) was used to produce figures, and MolScript (33) and Raster3D (34) were used to produce the figure. The figure was produced using the program GRASP (35).

The free R factor was calculated with the 5% of data omitted from the refinement. The R factor was calculated with the 5% of data omitted from the refinement. A bulk solvent correction (Babinet model with mask) was also used within the program REFMAC. See Table I for a summary of refinement statistics and model quality.

**RESULTS**

**Overall Structure of Native Maspin**—We have determined the crystal structures of native human maspin in two non-isomorphous forms, at 2.1- and 2.8-Å resolution, respectively (Table I and Fig. 1). The 2.8-Å form contains two molecules in the asymmetric unit (chains A and B) that are essentially identical (r.m.s.d. 0.2 Å over 365 Ca atoms). This low value reflects the non-crystallographic symmetry restraints employed during the refinement protocol. The high resolution 2.1-Å form also contains two molecules in the asymmetric unit; however, a superposition reveals that, although one molecule (chain A) is essentially identical to the two molecules in the 2.8-Å form, chain B contains significant differences in the position of the G-helix and the hairpin between s1B and s2B. These changes are discussed in detail below. Excluding these regions, the A and B chain of the 2.1-Å structure superpose with an r.m.s.d. of 0.5 Å over 339 Ca atoms. Both high and low resolution forms have similar unit cell parameters and related, but different, crystal packing arrangements. Although in both cases a crystallographic dimer is present within the asymmetric unit, the lack of shape complementarity and low buried surface area at the dimer interfaces suggests that this interaction is unlikely to be physiologically relevant. Moreover, maspin was purified in a monomeric form. Unless stated otherwise, structural analysis has been performed using chain A of the 2.1-Å resolution model.

**Analysis of Sequence Conservation**—The patterns of conservation within sub-branches of the intracellular (clade B) family of serpins (14) were investigated to infer residues uniquely important for maspin function. To account for limited species coverage and to ensure a comparable evolutionary timescale, only sequences arising from annioles were used. A structural alignment between maspin, ovalbumin, and PAI-2 (generated using Quanta) was used to seed a ClustalW (39) "profile" alignment of the remaining sequences. The analysis, in which amino acids with similar physico-chemical properties were treated as equivalent (A/V, M/I/L, F/Y, D/E, K/R, S/T, and N/Q), proceeded as follows: 1) conserved residues within the maspin subfamily were identified; 2) equivalent positions in other subfamilies (megsin, leukocyte elastase inhibitor, PAI-2, PI-6, bomaipin, and SCCA-1/-2), for which strict conservation was not evident, were tallied; and 3) a given alignment position was recorded where it was found to be absolutely conserved in maspin orthologues but not in more than half of the intracellular subfamilies. The short-listed residues therefore reflect positions putatively important to maspin function, but not to most other clade B serpins. None of these positions corresponded with the 50 most conserved of the serpin superfamily (14). Residues identified from the 2.1-Å structure as interacting with neighboring positions were removed; these amino acids were deemed less likely to participate in a ligand-protein binding event.

**RESULTS**

**Overall Structure of Native Maspin**—We have determined the crystal structures of native human maspin in two non-isomorphous forms, at 2.1- and 2.8-Å resolution, respectively (Table I and Fig. 1). The 2.8-Å form contains two molecules in the asymmetric unit (chains A and B) that are essentially identical (r.m.s.d. 0.2 Å over 365 Ca atoms). This low value reflects the non-crystallographic symmetry restraints employed during the refinement protocol. The high resolution 2.1-Å form also contains two molecules in the asymmetric unit; however, a superposition reveals that, although one molecule (chain A) is essentially identical to the two molecules in the 2.8-Å form, chain B contains significant differences in the position of the G-helix and the hairpin between s1B and s2B. These changes are discussed in detail below. Excluding these regions, the A and B chain of the 2.1-Å structure superpose with an r.m.s.d. of 0.5 Å over 339 Ca atoms. Both high and low resolution forms have similar unit cell parameters and related, but different, crystal packing arrangements. Although in both cases a crystallographic dimer is present within the asymmetric unit, the lack of shape complementarity and low buried surface area at the dimer interfaces suggests that this interaction is unlikely to be physiologically relevant. Moreover, maspin was purified in a monomeric form. Unless stated otherwise, structural analysis has been performed using chain A of the 2.1-Å resolution model.
Maspin adopts the typical native serpin fold: the structure comprises three antiparallel β-sheets (A, B, and C) surrounded by nine α-helices (hA–hI). The A β-sheet contains five β-strands, the N-terminal end of the RCL is located at the end of strand s5A, and the C-terminal end of the RCL joins strand s1C of the four-stranded C β-sheet. The six-stranded B β-sheet forms the hydrophobic core of the molecule. In contrast to many inhibitory serpins (for example, antithrombin and heparin cofactor II), the RCL is fully expelled from the A β-sheet (Fig. 1). Much of the RCL is highly mobile, as evidenced by poor electron density. Analysis by SDS-PAGE as well as mass spectrometry confirmed that the protein crystals contained only uncleaved material (data not shown). Maspin contains eight cysteine residues; however none of these form a disulfide bond nor do they lie in close proximity to one another. All buffers used during the production of recombinant maspin contained β-mercaptoethanol. As a result, in the higher resolution structure, derivatization of six out of the eight cysteine residues was apparent in the electron density of chain A, and three out of the eight in chain B.

Native Maspin Is in the Biologically Active Conformation—To confirm that the crystallized material was in the biologically active state, we assessed its ability to increase adhesion of cells to the extracellular matrix (8). As shown in Fig. 2A, pretreatment with maspin increased adhesion of human breast cancer cells to fibronectin by 20%. This is consistent with the 20–30% increase previously reported (8). In addition, native maspin and maspin cleaved within the RCL by cathepsin L have identical T_m values of 50 °C, as measured by CD spectroscopy. These data are consistent with published measurements of maspin stability and confirm that the molecule does not undergo the increase in thermostability upon cleavage that is characteristic of inhibitory serpins (Fig. 2B) (17, 40).

Structural Implications of Sequence Conservation—To identify important residues in maspin, we aligned all available maspin sequences (human, mouse, rat, chicken, and frog). Overall, these molecules display 43% sequence identity and 85% similarity with no obvious insertions or deletions (Fig. 3). Two of the sequences considered, that of chicken and frog maspin, have not been demonstrated to be bona fide orthologues. However, extensive sequence conservation with human maspin (>50% identity), the similarity in RCL length and sequence, as well as unusual features unique to the maspin structure (see below) support the inclusion of these sequences in the alignment. Throughout this report, maspin numbering is used, with antitrypsin numbering in parentheses.

Accessibility of Tyrosine Residues—It has been previously reported that maspin is phosphorylated on one or more tyrosine residues and that this modification may be important for an as yet uncharacterized role in signaling (41). Human maspin contains five tyrosine residues; with the exception of Tyr^{356} (372), these are conserved in the murine and rat sequences (Fig. 3). In contrast, many of these positions are substituted by phenylalanine or serine in the sequence of chicken and frog orthologues (Fig. 3). Four of these residues are buried; the exception, Tyr^{112} (138), is located at the top of the E-helix. The terminal hydroxyl group of Tyr^{112} is located at the bottom of a shallow solvent accessible cavity bounded by the A-, C-, and E-helices (Fig. 1). In this conformation there is insufficient space to accommodate a γ-phosphate moiety, and, furthermore, it is unlikely that any of the tyrosine residues, as observed in the native maspin structure, are sufficiently exposed to be accessible for modification by a kinase. This does not preclude the possibility of localized changes in solvent accessibility under appropriate circumstances. Although Tyr^{166}, located on strand 3A immediately
FIG. 3. Sequence alignment of all maspin-like proteins. The secondary structure of maspin is shown above the alignment, and numbering is for human maspin. Conserved residues are in bold and boxed. The positions of Lys$^{90}$ and Glu$^{115}$ are indicated by asterisks. Species are labeled as follows: hos, human; mmu, mouse; rno, rat; gga, chicken; xla, frog.
above the turn emerging from the F-helix, is buried in the maspin structure (accessible surface, 10 Å²), the portion of the F-helix near this residue is particularly flexible (42). Furthermore, in frog maspin this position is occupied by serine, a residue also capable of undergoing phosphorylation.

Comparison with Other Non-inhibitory Serpins—To date, the structure of only two other non-inhibitory serpins have been determined: pigment epithelium derived factor (PEDF) is a human extracellular angiogenesis inhibitor, and ovalbumin is an avian serpin of unknown function that belongs to the intracellular branch of the family (40, 43). Recently, a 3.1-Å x-ray crystal structure of a maspin mutant in which all cysteine residues were mutated to alanine has been reported (22).

Maspin has a similar number of hydrogen bonds but fewer salt bridges (412/104) in comparison to PEDF (374/168) and ovalbumin (412/206). Consistent with these data both PEDF and ovalbumin have a higher T_m (58 °C and 70 °C (44, 45)) than maspin (50 °C (Ref. 17 and this study)). Maspin and PEDF share 25% sequence identity, and a structural comparison reveals that these molecules superpose poorly, with an r.m.s.d. of 1.3 Å/atom over 248 C atoms (Fig. 4). Major structural changes (>1.5 Å) are observed in the A-, D-, E-, F-, G-, and H-helices as well as s1A and s2A, at the bottom of the A β-sheet (Fig. 4). Maspin and ovalbumin share 31% identity and superpose with an r.m.s.d. of 1.5 Å/atom over 335 C atoms (Fig. 4). Again, major structural changes are observed in the A-, D-, E-, F-, G-, and H-helices as well as the top of s1A and the base of s2A (Fig. 4). The comparison reveals that the N-terminal portion of the RCL is in a similar conformation in all three serpins, tightly interacting with the body of the molecule (Figs. 1 and 4). Structural studies to date reveal that the RCL of many inhibitory serpins is either partially inserted, or shifted to a position that facilitates insertion into the A β-sheet.

Interestingly, the surface electrostatic potential of maspin reveals an asymmetric charge distribution similar to that of another angiogenesis-inhibiting serpin, PEDF (43). One face of maspin, centered on the D-helix, includes a cluster of basic residues (Fig. 5A), whereas the opposite face is predominantly negatively charged (Fig. 5B). By analogy with PEDF, this negatively charged region could represent a possible collagen binding site (21).

Maspin Contains a Buried Salt Bridge on the Periphery of the Shutter—The shutter is a conserved region located in the center of the serpin fold that is important for controlling conformational change in inhibitory molecules (Fig. 1) (46). Many non-inhibitory serpins contain non-conservative substitutions in this region, presumably because these molecules have lost the requirement to undergo the S to R transition. One unusual feature of the maspin structure is a buried salt bridge formed between Lys⁹⁰ (116) on s2A and Glu¹¹⁵ (141) on s1A (Fig. 5C).

Two water molecules are also coordinated into this region and help dissipate the buried charge. Analysis of an extensive sequence alignment of the serpins as well as available serpin structures reveals that the salt bridge is predicted to be present only in maspin orthologues (Fig. 3) (14). The majority of serpins contain an asparagine or serine residue, at the position equivalent to Lys⁹⁰ (116), which forms a hydrogen bond to the side chain of the highly conserved shutter region residue Asn¹⁶³ (186) (Fig. 5D). In maspin, the side chain of Asn¹⁶³ instead forms a hydrogen bond to the side chain of Ser¹⁸⁶ (60) on the B-helix as well as to His¹²⁰ (334) on s5A.

The Lys⁹⁰/Glu¹¹⁵ salt bridge lies at the base of a deep solvent-accessible cavity (45 Å³) centered on the shutter region and formed by residues from the D-helix (Phe⁷⁰, Val⁷³, Thr⁷⁴, and Val¹⁷¹), the B-helix (Ile¹³⁸) and s2A (Leu¹⁸⁸) (Figs. 1, 5A, and 5C). In comparison to ovalbumin, PEDF, and other structurally characterized serpins, the interaction causes an unusual distortion in s1A that results in a significant repositioning of Lys¹¹⁴ (140) (Fig. 5C). Although Lys¹¹⁴ is not conserved in all maspin-like proteins, in human maspin this residue contributes to a cluster of positively charged residues (Lys⁹⁰ (105), Lys⁹⁷ (113), Arg¹¹⁰ (117), Lys¹⁰⁹ (135), and Arg¹¹⁶ (136)) (Figs. 1, 5C, and 5D) that are conserved in all maspin orthologues (Fig. 3). Interestingly, this cluster of charged residues is centered on the D- and E-helices, a region that is responsible for cofactor binding in many serpins.

Conformational Heterogeneity at the G-helix—A structural superposition of all the maspin chains determined in this study (2.1 and 2.8 Å), and the previously determined, non-isomorphous 3.1-Å structure 1XQG (22), reveals striking conformational heterogeneity at and around the G-helix (Figs. 5B, 5D, and 6A). All five chains (2.1-Å A/B, 2.8-Å chains A/B, and 1XQG chain A) can be superposed with an r.m.s.d. of 0.56 Å over 339 Ca atoms and show a high degree of similarity in all regions apart from the G helix and the hairpin connecting strands s1B and s2B.

Two conformational variants of maspin are presented in this study. Chain B of the 2.1-Å structure, which we have termed the “closed” conformation, corresponds in morphology with the 3.1-Å structure (1XQG), as well as the packing exhibited by other structurally characterized serpins: the G-helix makes close contacts with the s2C-s4B and s1B-s2B hairpin loops and the A-helix (Fig. 5E, 6A, and 6B). In contrast, both chains of the 2.8-Å structure and chain A of the 2.1-Å structure correspond
with a novel “open” serpin conformation identified for the first time here. In the open conformation, the interfaces at opposite poles of the G-helix are widened: eight water molecules are buried between the G-helix and the s1B–s2B hairpin, compared with just one in the closed form. Accordingly, the open form also contains three additional -mercaptoethanol-derivatized cysteine residues, indicating the increased degree of access to the interfaces made with the G-helix. The shift at the G-helix that gives rise to these two distinct conformations can be described as a rotation of 5° of the G-helix that hinges around residue 249 (Figs. 5E and 6A). This produces shifts in excess of 3 Å; for example, Glu246 moves by 3.7 Å. To accommodate this movement in the G-helix, the hairpin loop connecting strands s1B and s2B shifts by ∼2 Å. Furthermore, this conformational change gives rise to large differences in packing between the G-helix and the underlying B-sheet and the A helix.

The open and closed forms differ in their surface electrostatic nature (Fig. 5B), such that the continuous band of negative charge spanning the surface of the G-helix in the open form is altered in the alternative conformation. This change is primarily a result of Glu244 and Glu247 moving from a buried to an exposed position.

Residues Uniquely Conserved in Maspin Partition to Opposite Sides of the Molecule—Thirty-two amino acid positions (9%) are conserved among maspin orthologues but not in other clade B subfamilies (Fig. 7). Most (28) of these residues are surface-exposed (>20 Å²), suggesting that they would be available to mediate higher order interactions with a ligand or protein. Structural analysis reveals that one of these residues is involved in non-main-chain hydrophobic contacts and two others form side-chain-mediated hydrogen bonds. Specifically, Leu80 contributes to a hydrophobic patch that stabilizes the final turn of the D-helix. In ovalbumin, this region adopts an extended conformation as a result of a proline residue substituted at this position, whereas the top of the D-helix is disordered in PAI-2. His59 forms a (pH-dependent) salt bridge with Asp14, linking the start of the C-helix to the A-helix; in contrast this position is arginine in ovalbumin and glutamine in human PAI-2. Ser107 on the E-helix forms a hydrogen bond with Glu103 (also on the E-helix). In ovalbumin Ser107 is substituted by a cysteine, which forms a disulfide bond to the B-helix.

Of the remaining residues that do not appear to play structural roles, 10 are acidic, 4 basic, 8 polar, and 7 non-polar. These residues largely show a distinct partition into two groups: a predominantly acidic patch extending from the s4C–s3C loop and G-helix down to the I-helix; and a mixed basic/polar stretch in the loops abutting the D-helix and s2C/s3C. As detailed previously, the acidic patch that surrounds the G-helix...
is implicated in an altered negative charge distribution during the shift between the open and closed forms of maspin.

**DISCUSSION**

The x-ray crystal structures of maspin presented in this work reveal for the first time that the G-helix of maspin is capable of undergoing a significant conformational change. Furthermore, this is the first time a serpin has been demonstrated to undergo movement in this region of the molecule. The two structures, reflecting maspin in the open and closed conformation, show that rotation of the G-helix alters the local charge distribution, suggesting that this movement represents a conformational “switch.” Furthermore, sequence analysis reveals that the broader patch of acidic residues within which the G-helix is localized is a highly conserved feature of maspin orthologues, and unique with respect to other intracellular serpins. Taken together, these data suggest that the face of the maspin molecule incorporating the s4C–s3C loop, G-helix, and I-helix comprises an important aspect of maspin function.

Overall, the maspin molecule corresponds in topology with the native serpin fold. The material crystallized is biologically active, and biophysical data confirm that native maspin is unable to undergo the S to R transition upon cleavage within the RCL. Previous work has demonstrated that the RCL sequence of maspin is important for function; in particular, mutation of the Arg340 to Ala abolishes the ability of the molecule to prevent cell adhesion (8). Much of the RCL is disordered in the high resolution crystal structure, and no interactions between the Arg340 and the body of the serpin are observed. Based upon these data, it is suggested that the RCL mediates interactions with an uncharacterized target, rather than through the formation of intramolecular contacts.

The location of cysteine residues in maspin is inconsistent with intramolecular disulfide bonds forming in the extracellular milieu. Furthermore, the structure reveals that none of the five tyrosine residues in maspin are sufficiently exposed to allow phosphorylation by a kinase. However, we cannot exclude the possibility that, for example, interaction with a binding partner may promote subtle conformational shifts that expose one or more tyrosine residues. We note that Tyr112 (at the top of the E-helix), Tyr246 (in the loop between hD and s2A), and Tyr266 are located in regions of the molecule that, in other serpins, are capable of undergoing substantial conformational change. Such a rearrangement would not be unprecedented even in non-inhibitory serpins: for example, the x-ray crystal structure of S ovalbumin reveals a conformational change in the region around Tyr122 (s1A/s2A and the E-helix (47)).

Analysis of the structure reveals a buried salt bridge is formed between Lys390 on s2A and Glu115 on s1A. Such interactions are extremely rare and generally destabilize proteins owing to the energetic expense of placing charged residues in the hydrophobic core (48, 49). Previous studies have suggested that one general role of such an interaction may be in determining specificity or geometry in protein-cofactor interactions (48, 49). Comparison with other serpins reveals that, in maspin, the side chain of Glu115 is shifted toward the body of the molecule to neutralize the positive charge of Lys390. This salt bridge causes a prominent bulge at the N-terminal end of s1A and reveals a cavity beneath the D-helix; it is suggested that this region, as in other serpins (for example, antithrombin and plasminogen activator inhibitor-1), may be important for interaction with a binding partner. Furthermore, the distortion in secondary structure caused by the salt bridge introduces Lys114 into the center of a cluster of conserved positively charged residues. Maspin is able to bind heparin (7), and it is possible that these residues represent a heparin binding site (50).

Comparison of the three independently derived crystal forms of maspin reveals a major conformational rearrangement in the region around the G-helix. This region of the serpin scaffold was previously thought to be conformationally inert (51, 52); a structural comparison of several serpin pairs does not reveal conformational mobility around the G-helix (38, 53). The switch from the closed to the open form involves the G-helix rotating by approx 5° and translating ~3 Å on the underlying B-sheet, with corresponding 2-Å shifts in the hairpin connecting strands s1B and s2B of the B-sheet. Strikingly, our data thus reveal that maspin is able to undergo a novel conformational change in a region of the serpin scaffold never before associated with conformational mobility. Furthermore, the rearrangement results in Glu244 and Glu247 moving from a partially buried to an exposed position. As a result, maspin displays an elongated negatively charged patch, which can be modulated by conformational switching at the G-helix (Fig. 5, B and E). Significant electrostatic patches can be indicative of regions that mediate protein-protein and/or ligand interactions (54). Although many serpins bind cofactors on or around the D-helix, a negatively
charged region around the G-helix has been implicated in collagen binding in the non-inhibitory serpin PEDF (55). Previous studies have shown that maspin is able to bind collagen (21), and it is suggested that the conformational flexibility in this region in maspin highlights a functionally important region of the molecule, most likely important for binding cofactors such as collagen.

In conclusion, the work provides a high resolution structural foundation that opens up new avenues for research aimed at understanding the molecular basis of maspin function. The structure reveals novel features in the region around the D-helix and G-helix that may prove useful in the design of specific therapeutics aimed at modulating or emulating maspin function.

Acknowledgments—We thank the staff at BioCARS and the Australian Synchrotron Research Program for assistance.

REFERENCES

1. Zou, Z., Anisowicz, A., Hendrix, M. J., Thor, A., Neveu, M., Sheng, S., Rafal, K., Seftor, E., and Sager, R. (1994) Science 263, 526–529
2. Zhang, M. (2004) Prost. Biol. 9, 2218–2226
3. Silverman, G. A., Bird, P. L., Carr, R. W., Church, F. C., Coughlin, P. B., Gettins, P. G., Irving, J. A., Lomas, D. A., Luke, C. J., Moyer, R. W., Pemberton, P. A., Remond-O’Dell, E., Salvesen, G. S., Travis, J., and Whistock, J. C. (2001) J. Biol. Chem. 276, 33293–33296
4. Reis-Filho, J. S., Milanezi, F., Silva, P., and Schmitt, F. C. (2001) Oncogene 20, 4078–4083
5. Jiang, N., Meng, Y., Zhang, S., Mensah-Osman, E., and Sheng, S. (2002) Biochemistry 41, 4989–4998
6. Sheng, S., Carey, J., Seftor, E. A., Dias, L., Hendrix, M. J., and Sager, R. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 11669–11674
7. Ngamkitidechakul, C., Marqueze, D. J., Bourge, J. M., and Sager, R. (1995) Biochem. J. 317, 3186–3190
8. Gao, F., Shi, H. Y., Daughtey, C., and Zang, M. (2004) Development 131, 1417–1437
9. Zhang, M., Volpert, O., Shi, Y. H., and Beauc, S. (2000) Biochem. J. 356, 196–199
10. Zou, Z., Cho, G., Nakaia, A. K., Connell, T., Saito, S., Moxon, R., and Sager, R. (2002) J. Biol. Chem. 277, 6019–6024
11. Kim, S., Han, J., Kiton, M., and Park, C. (2004) Cancer Res. 64, 6090–6095
12. Whistock, J., Skinner, R., and Lesk, A. M. (1998) Trends Biochem. Sci. 23, 63–67
13. Irving, J. A., Pike, R. N., Less, A. M., and Whistock, J. C. (2000) Genome Res. 10, 1845–1854
14. Huntington, J. A., Cappai, R., Fenn, S. C., Henry, A., McKinnon, W. J., Galatis, D., Hesse, L., Mulhaupt, G., Beyreuther, K., Masters, C. L., and Parker, M. W. (1999) Nat. Struct. Biol. 6, 327–331
15. James, E. L., Whistock, J. C., Gore, M. G., and Bottomley, S. P. (1999) J. Biol. Chem. 274, 5432–5448
16. Simonovic, M., Gettins, P. G., and Velz, K. (2001) Proc. Natl. Acad. Sci. U. S. A. 98, 12878–12883
17. Osawa, M., Kikkawa, K., and Kikkawa, K. (2002) J. Biol. Chem. 277, 8149–8157
18. Post, D. W., and Bottomley, S. P. (2001) J. Biol. Chem. 272, 32725–32731
19. Whistock, J. C., Skinner, R., and Lesk, A. M. (2000) J. Mol. Biol. 295, 685–699
20. Thompson, J. D., Higgs, D. G., and Gibson, T. J. (1994) Nucleic Acids Res. 22, 4675–4680
21. Stein, P. E., Leslie, A. G., Finch, J. T., Turnbull, W. G., McLaughlin, P. J., and Carrell, R. W. (1990) Nature 347, 99–102
22. Ossenbrink, R. A., Khalkhali-Elli, Z., Schneider, G. B., Seftor, E. A., Seftor, R. E., Koland, J. G., and Hendrix, M. J. (2002) Biochem. Biophys. Res. Commun. 295, 800–805
23. Hendsch, Z. S., and Tidor, B. (1994) Protein Sci. 3, 211–226
24. Nakamura, H. (1996) Q. Rev. Biophys. 29, 1–90
25. Rossjohn, J., Cappai, R., Fenn, S. C., Henry, A., McKinnon, W. J., Galatis, D., Hesse, L., Mulhaupt, G., Beyreuther, K., Masters, C. L., and Parker, M. W. (1999) J. Mol. Biol. 295, 327–331
26. Whistock, W. D., and Bottomley, S. P. (2001) J. Biol. Chem. 276, 6382–6387
27. Whistock, J. C., and Whistock, R. W. (1995) Nat. Struct. Biol. 2, 96–113
28. Yamashita, K., Takahashi, N., and Hirose, M. (2005) J. Biol. Chem. 278, 35524–35530
29. Al-Ayyoubi, M., Gettins, P. G., and Velz, K. (2004) J. Biol. Chem. 279, 55540–55544
30. Kapust, R. B., Tozzer, J., Fox, J. D., Anderson, D. E., Cherry, S., Copeland, T. D., and Waugh, D. S. (2003) Protein Eng. 16, 993–1000
31. Simonovic, M., Gettins, P. G., and Volz, K. (2004) J. Biol. Chem. 279, 45400–45407
32. Kraf, P. J. (1991) Acta Crystallogr. A 47, 110–119
33. Murshudov, G. N., Deacon, A. D., and Dodson, E. J. (1997) Acta Crystallogr. D 53, 240–250
34. Navaza, J. (1993) Acta Crystallogr. D Biol. Crystallogr. 59, 1367–1372
35. Rodriguez, R., Chinea, G., Lopez, N., Pons, T., and Vriend, G. (1998) Bioinformatics 14, 523–528
36. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Acta Crystallogr. D 53, 240–250
37. Storoni, L. C., McCoy, A. J., and Read, R. J. (2004) Acta Crystallogr. D Biol. Crystallogr. 60, 432–438
38. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Acta Crystallogr. D Biol. Crystallogr. 54, 905–921
39. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard (1991) Acta Crystallogr. A 47, 110–119
40. Merritt, E. A. (1994) Acta Crystallogr. D Biol. Crystallogr. 50, 869–873
41. Nicholls, A. (1992) GRASP, Graphical Representation and Analysis of Surface Properties, New York
The High Resolution Crystal Structure of the Human Tumor Suppressor Maspin Reveals a Novel Conformational Switch in the G-helix
Ruby H. P. Law, James A. Irving, Ashley M. Buckle, Katya Ruzyla, Marguerite Buzza, Tanya A. Bashtannyk-Puhalovich, Travis C. Beddoe, Kim Nguyen, D. Margaret Worrall, Stephen P. Bottomley, Phillip I. Bird, Jamie Rossjohn and James C. Whisstock

J. Biol. Chem. 2005, 280:22356-22364.
doi: 10.1074/jbc.M412043200 originally published online March 10, 2005

Access the most updated version of this article at doi: 10.1074/jbc.M412043200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 52 references, 19 of which can be accessed free at http://www.jbc.org/content/280/23/22356.full.html#ref-list-1