Kotler, M., W. Danho, R.A. Katz, J. Leis, and A.M. Skalka. 1988. Synthetic peptides as substrates and inhibitors of a retroviral protease. *Proc. Natl. Acad. Sci.* **85:** 4185.

———. 1989. Avian retroviral protease and cellular aspartic proteases are distinguished by activities in peptide substrates. *J. Biol. Chem.* **264:** 3428.

Miller, R. 1987. Proteolytic self-cleavage of hepatitis B virus core protein may generate serum e antigen. *Science* **236:** 722.

Miller, M., M. Jaskólski, J.K.M. Rao, J. Leis, and A. Wlodawer. 1989. Crystal structure of a retroviral protease proves relationship to aspartic protease family. *Nature* **337:** 576.

Navia, M.A., P.M.D. Fitzgerald, B.M. McKeever, C.-T. Leu, J.C. Heimbach, W.K. Herber, I.S. Sigal, P.L. Darke, and J.P. Springer. 1989. Three-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1. *Nature* **337:** 615.

Schneider, J. and S. Kent. 1988. Enzymatic activity of a synthetic 99 residue protein corresponding to the putative HIV-1 protease. *Cell* **54:** 363.

Seelmeier, S., H. Schmidt, V. Turk, and K. van der Helm. 1988. Human immunodeficiency virus has an aspartic-type protease that can be inhibited by pepstatin A. *Proc. Natl. Acad. Sci.* **85:** 6612.

Toh, H., M. Ono, K. Saigo, and T. Miyata. 1985. Retroviral protease-like sequence in the yeast transposon Ty1. *Nature* **315:** 691.

Wang, B.-C. 1985. Resolution of phase ambiguity in macromolecular crystallography. *Methods Enzymol.* **115:** 90.

Crystal Structure of a Retroviral Proteinase from Avian Myeloblastosis Associated Virus

S.I. Foundling,[1] F.R. Salemme,[1] B. Korant,[1] J.J. Wendoloski,[1] P.C. Weber,[1] A.C. Treharne,[1] M.C. Schadt,[1] M. Jaskólski,[2] M. Miller,[2] A. Wlodawer,[2] P. Strop,[3] V. Kostka,[3] J. Sedlacek,[3] and D.H. Ohlendorf[1]

[1]Central Research and Development Department
Du Pont Experimental Station, Wilmington, Delaware 19880-0228

[2]Crystallography Laboratory, NCI-Frederick Cancer Research Facility, BRI-Basic Research Program, Frederick, Maryland 21701

[3]Institutes of Organic, Biochemistry, and Molecular Genetics
Czechoslovak Academy of Science, 16610 Prague 6, Czechoslovakia

Crystallographic and structural analyses are reported for the recombinant retroviral proteinase MAV, derived from the myeloblastosis-associated virus. A complete X-ray data set to 2.25 Å resolution was collected from a single crystal using a position-sensitive detector. Data reduction and analysis show that the proteinase crystallizes in the trigonal space group $P3_121$ with unit cell dimensions, $a = b = 88.9$ Å, $c = 78.8$ Å. The proteinase exists as a dimer in the crystallographic asymmetric unit and shares an overall structural similarity to the aspartic proteinase family of enzymes.

The retroviral genome incorporates three open reading frames that code for the *gag*, *pol*, and *env* polyproteins. These polyproteins are cleaved during viral maturation into functional enzymes and capsid proteins through the action of a specific retroviral proteinase (Moelling et al. 1980; Dickson et al. 1984; Katoh et al. 1985). In the human immunodeficiency virus, type 1 (HIV-1), the *gag-pol* polyprotein comprises not only the structural *gag* proteins, but also three enzymes, the p10 PR proteinase, p66 (p51) RT reverse transcriptase, and p32 IN integrase (Ratner et al. 1985). In the avian sarcoma/leukosis retroviruses, the equivalent proteinase is located in the *gag* reading frame (Fig. 1) and is somewhat larger in size than its

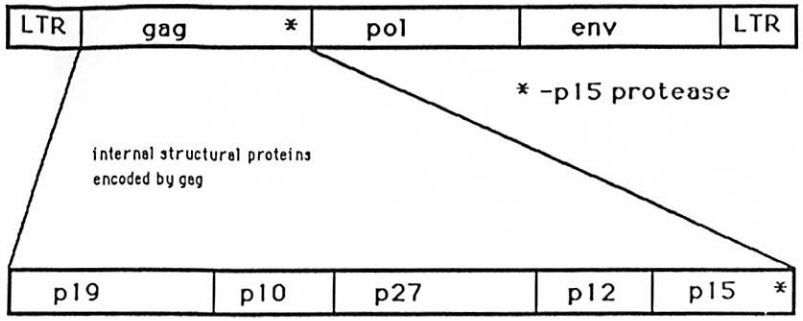


Figure 1 MAV proteinase. Avian sarcoma/leukosis retroviruses incorporate a p15 maturation proteinase in the ORF-1 (*gag*) reading frame. In MAV, the proteinase has 124 amino acid residues and a molecular weight of 13,500.

counterparts in the human retroviruses (Toh et al. 1985; Pearl and Taylor 1987). Because of their critical role in viral replication, maturation proteinases are prime targets for structural studies aimed at the rational design of antiviral compounds.

METHODS

Protein was isolated from a recombinant *Escherichia coli* system that expressed an artificial precursor incorporating the proteinase as a fused gene product. Autolytically liberated proteinase was purified using standard chromatographic techniques (Sedlacek et al. 1988) and characterized by SDS-PAGE and isoelectric focusing to be homogeneous prior to crystallization. Following dialysis and concentration by ultrafiltration, the protein was set up in hanging drop experiments to establish crystallization conditions for X-ray analysis (Cox and Weber 1987, 1988). The best crystals grew as hexagonal rods (0.1 mm–0.4 mm) from 8-μl drops (Fig. 2) prepared by mixing 4 μl of 0.1 mM protein with 4 μl of reservoir solution (0.05 M citrate/phosphate [pH 5.6] 8–12% $[NH_4]_2SO_4$, 5% dimethyl formamide). The reservoir additionally contains 0.05% β-mercaptoethanol. Crystals typically grow within 1 week at room temperature.

Native data sets were collected from a number of crystals using a Nicolet/Xentronics imaging proportional counter with monochromated CuKα radiation generated by an Elliot GX-21 rotating anode X-ray source. The data used in the present analysis included 2845 0.25-degree frames collected from a single

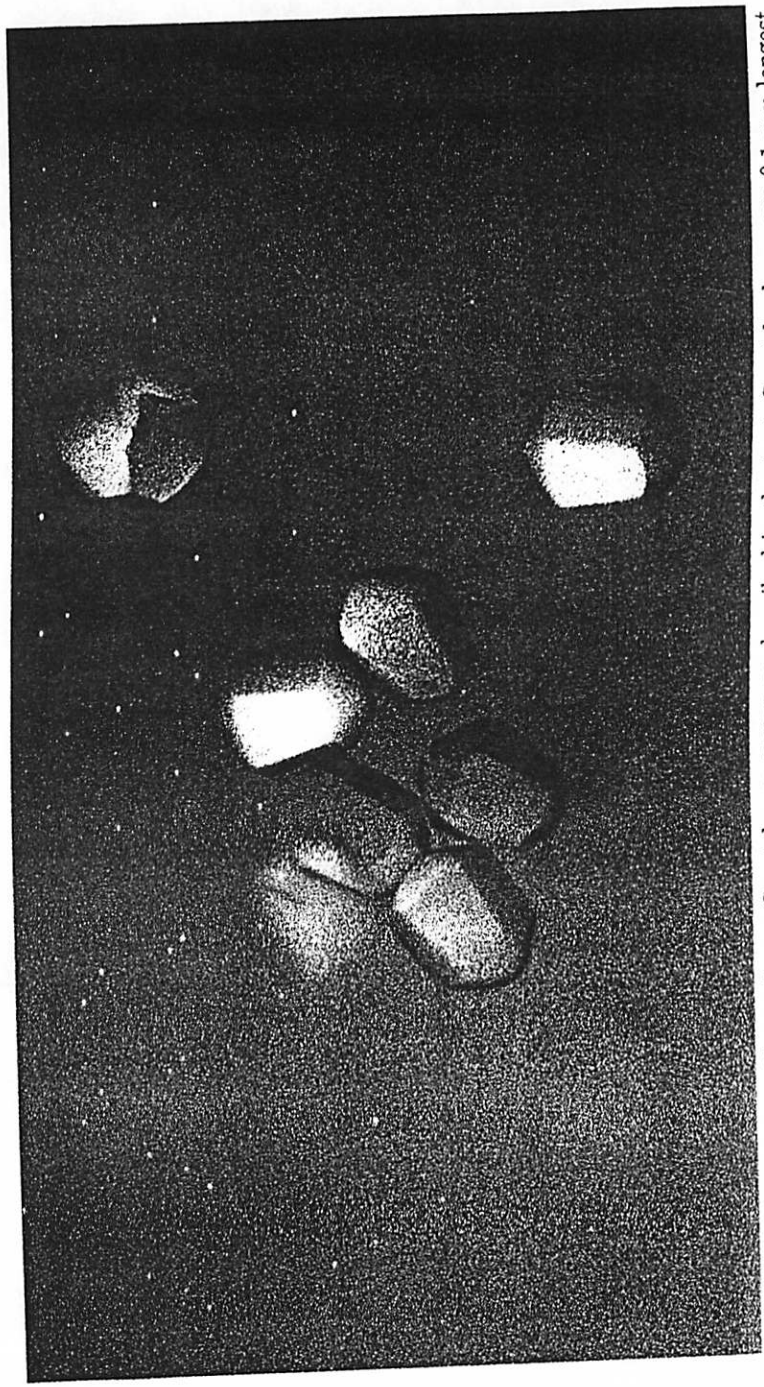


Figure 2 Crystals of MAV proteinase. Crystals are grown as described in the text. Crystals shown are 0.1 mm longest dimension.

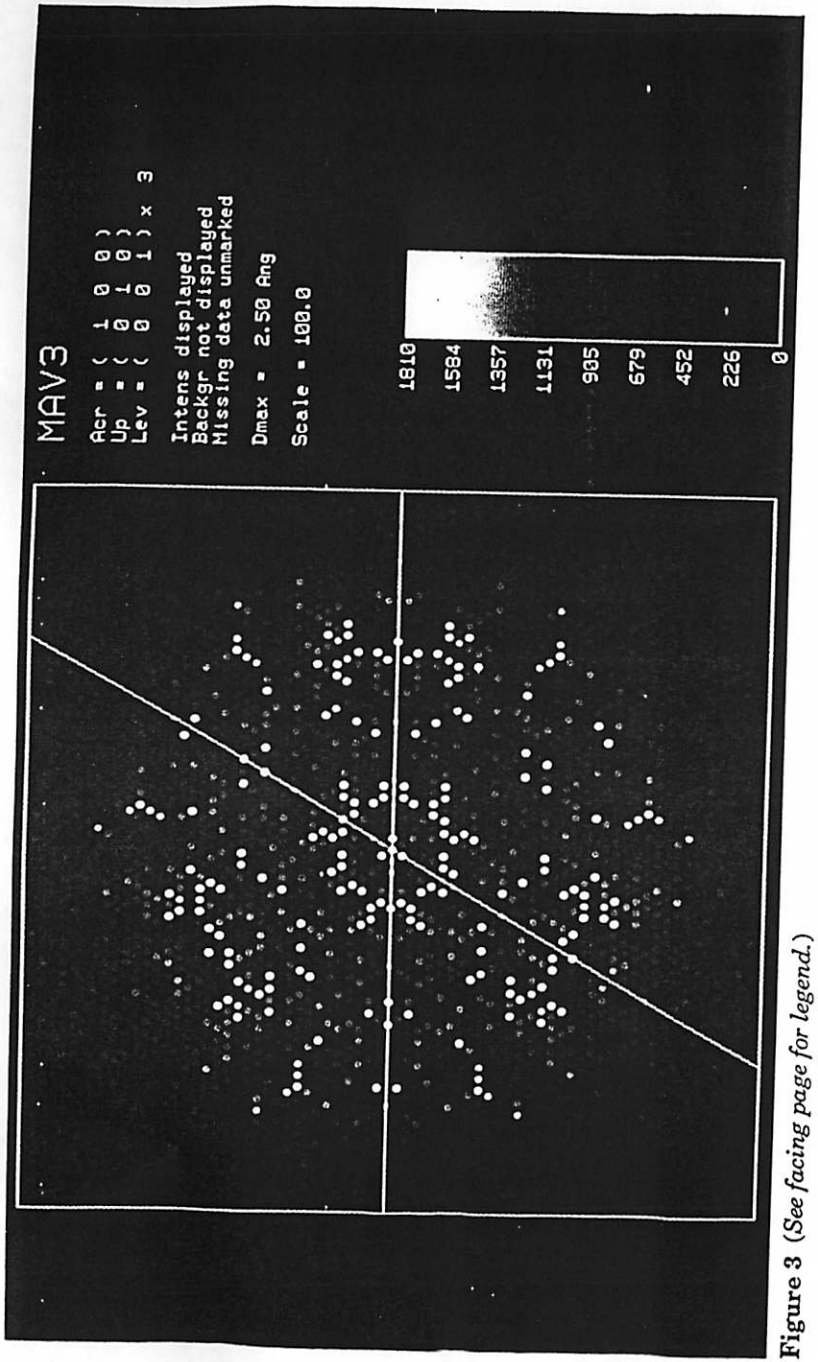


Figure 3 *(See facing page for legend.)*

crystal. Data collection continued for 12 days, during which there was little appreciable radiation damage to the crystal. The diffraction data were then indexed and integrated using the XENGEN software (Howard et al. 1987). The MAV proteinase crystallizes in a trigonal unit cell with dimensions $a = b = 88.9$ Å, $c = 78.8$ Å, $\alpha = \beta = 90°$, $\gamma = 120°$. Subsequent examination of the indexed diffraction pattern determined the space group as $P3_121$ (Fig. 3). The data set is virtually complete to 2.25 Å with 196,542 observations of 16,968 unique reflections (4 missing). Intensities were well measured to 2.4 Å where the mean intensity was 2σ above background, dropping to 1σ above background at 2.25 Å resolution. The unique reflections merged with a σ-weighted R_{sym} on intensities of 8.6%.

The MAV structure was solved by molecular replacement using a partially refined Rous sarcoma virus (RSV) proteinase model (Miller et al. 1989). The results of six-dimensional searches (Fujinaga and Read 1987) of this model against the MAV reflection data in the resolution ranges of 5 Å–4 Å and 3.5 Å–3.2 Å gave solution peaks with correlation coefficients of 0.737 and 0.697, respectively. The solution corresponded to only a minor shift of the input RSV molecule. From this model, initial electron density maps were calculated at 3.2 Å. The R-factor for this model was 0.34.

RESULTS

The MAV proteinase crystallizes with the native dimer in the crystallographic asymmetric unit, as is observed for the essentially isomorphous RSV proteinase (Miller et al. 1989). The monomers of the dimer are related to each other by a non-crystallographic dyad axis, so that the two active-site aspartic acids are brought into close juxtaposition at the base of the substrate-binding cleft. In this respect, the topology of the MAV dimer is similar to that observed in the bilobal aspartic proteinases, of which penicillopepsin and endothiapepsin are examples (James and Sielecki 1983; Blundell et al. 1985). Each monomer is composed of two twisted β sheets of four β strands, which form the walls of an irregular barrel. Although the MAV structure is not fully refined, Figure 4, which shows a section of

Figure 3 (hk3) diffraction zone synthesized from intensities of integrated reflections. Lines show the directions of the a^* and b^* axes. Sides of box are at 2.5 Å resolution.

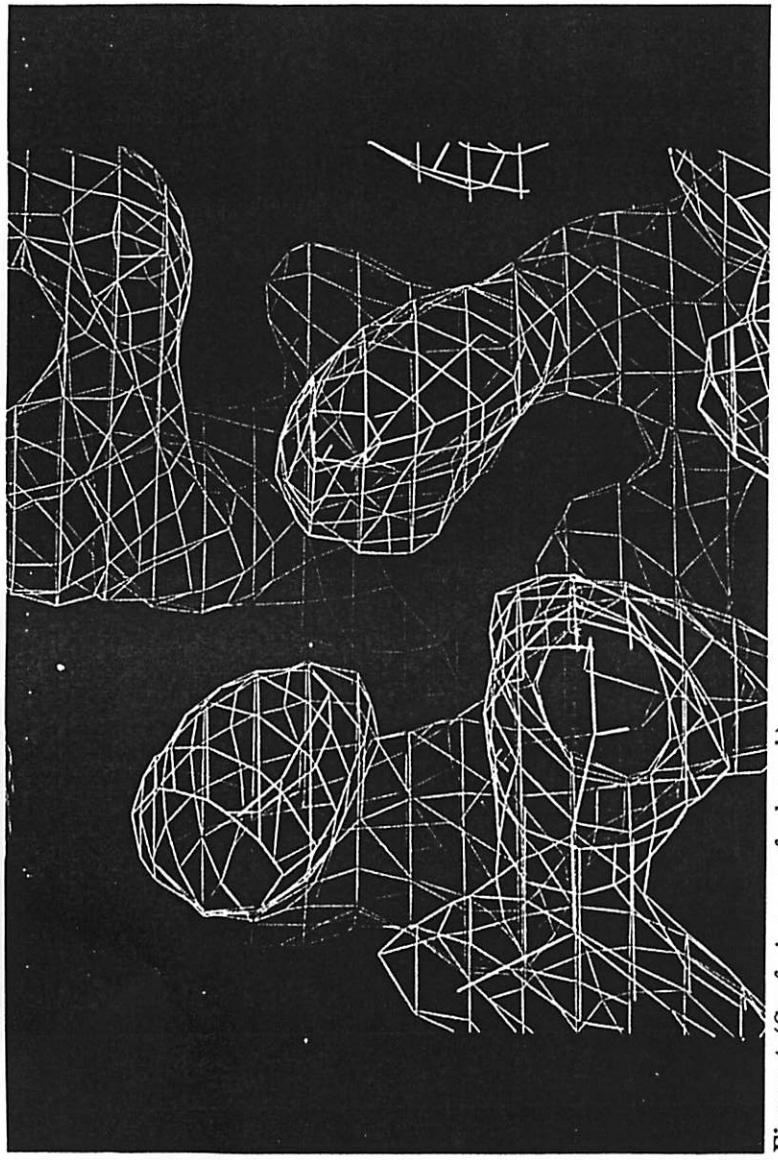


Figure 4 *(See facing page for legend.)*

the $2F_o-F_c$ electron density incorporating the active-site Asp*-Ser-Gly segments, illustrates the quality of the current X-ray map.

Refinement of the MAV model is continuing, with each monomer of the dimer being refined independently (Hendrickson 1980). This will allow any local structural differences that may exist between monomers as a result of the different molecular environments to be preserved. Remodeling into electron density maps using FRODO (Jones 1978) is aided with a program that uses a crystallographic structural data base for checking side-chain rotamers and secondary structural motifs (Weber et al. 1989). The current model incorporates neither solvent molecules nor individual temperature factors and gives an R-factor of 0.28 at 2.4 Å.

DISCUSSION

Comparison of the sequences of retroviral proteinases demonstrates the high degree of homology shared among these molecules (Tables 1 and 2). The active-site aspartic acid is indicated with an asterisk, and the highly conserved sequences common across all retroviral proteinases are indicated with points. The conserved sequences form a three-strand β sheet that structurally resembles the greek letter ψ. Two copies of this motif are found in the bilobal aspartic proteinases. The conserved Gly-Arg-Asp/Asn residues are positioned at the top of a common helix in MAV and RSV. This helix is apparently absent in the HIV-1 structure (Navia et al. 1989). The degree of sequence homology suggests that the structures of the MAV and RSV maturation proteinases can be used to model the structures of other retroviral proteinases. Given the structural homology seen between these retroviral proteinases and the aspartic proteinases, it should be possible to extrapolate features of inhibitor binding observed in the latter to models for inhibitors bound to retroviral proteinases, including HIV-1 proteinase. From these models, it may be possible to rationally design proteinase inhibitors and so block viral maturation and infectivity of the virus. Crystallographic analyses of inhibitors bound to MAV proteinase will provide a better basis with which

Figure 4 The $2F_o-F_c$ map for the catalytic aspartic acid and adjacent residues from each monomer. The view is toward the bottom of the active-site cleft.

Table 1 Sequence Alignments of MAV, RSV, HIV-1, and HIV-2 Retroviral Proteinases

```
                    1        10        20        30       *40        50
MAV                 LAMTMEHKDRPLVRVILTNTGSHPVKQRSVYITALLDSGADITIISEEDW
RSV                 LAMTMEHKDRPLVRVILTNTGSHPVKQRSVYITALLDSGADITIISEEDW
HIV-1               PQITLW--QRPLVTIKIG----------GQLKEALLDTGADDTVLEEMSL
HIV-2               PQFSLW--KRPVVTAYIE----------GQPVEVLLDTGADDSIVAGIEL
                    h hth     RPhV h h             h  hLLDtGAD thh    h
                      s                                    s   s

                           60        70        80        90        100
MAV                 PADRPVMEAANPQIHGIGGGIPMRKSRDMIEVGVINRDGSLERPLLLFPA
RSV                 PTDWPVMEAANPQIHGIGGGIPMRKSRDMIELGVINRDGSLERPLLLFPA
HIV-1               PGKWKPKM-----IGGIGGFIKVRQY-DQILIEIC----GHK-AIGT-VL
HIV-2               GNNYSPIK-----VGGIGGFINTKEY-KNVEIEVL----NKK-VRAT-MT
                                GIGG I  +     h h h              h

                          110       120
MAV                 VAMVRGSILGRDCLQGLGLRLTNL
RSV                 VAMVRGSILGRDCLQGLGLRLTNL
HIV-1               VGPTPVNIIGRNLLTQIGCTLNF-
HIV-2               TGDTPINIFGRNILTALGMSLNL-
                          IhGR hL  hGh L
```

The numbering is that adopted for the primary sequence of MAV (RSV), with the single letter amino acid code defining residue identity. Active-site Asps are identified with an asterisk, and the highly conserved sequences are identified with points. Below the sequence alignment are positions in the sequence that are conserved for structural reasons, either buried hydrophobic (h) or residues that are conserved at turns in the structure.

Table 2 Retroviral Proteinase Sequence Homology

	HIV-1	HIV-2	MAV	RSV
HIV-1		70%	42%	41%
HIV-2			42%	43%
MAV				99%
RSV				

The homology table was constructed using a sequence analysis software package, WISCONSIN (Devereux and Haeberli 1985).

to guide these modeling studies. Such analyses are under way and are the goal of structural studies on these retrovirally encoded enzymes.

ACKNOWLEDGMENTS

This research was sponsored in part by the National Cancer Institute, the Department of Health and Human Services (DHHS), with Bionetics Research, Inc. (BRI) under contract NO1-CO-74101.

REFERENCES

Blundell, T.L., J.A. Jenkins, L.H. Pearl, B.T. Sewell, and V.B. Pedersen. 1985. The high resolution structure of endothiapepsin. In *Aspartic proteinases and their inhibitors* (ed. V. Kostka), p. 151. de Gruyter, Berlin.

Cox, M.J. and P.C. Weber 1987. Experiments in automated protein crystallization. *J. Appl. Crystallogr.* 20: 366.

———. 1988. An investigation of protein crystallization parameters using successive automated grid searches. *J. Appl. Crystal Growth* 90: 318.

Devereux, J. and P. Haeberli. 1985. WISCONSIN sequence analysis software package. University of Wisconsin Biotechnology Centre, Madison, Wisconsin (WS 53705).

Dickson, C., R. Eisenman, H. Fan, E. Hunter, and N. Teich. 1984. Protein biosynthesis and assembly. In *Molecular biology of tumor viruses*, 2nd edition: *RNA tumor viruses*, 1/Text (ed. R. Weiss et al.), p. 513. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Fujinaga, M. and R. Read. 1987. Experiences with a new translation-function program. *J. Appl. Crystallogr.* 20: 517.

Hendrickson, W.A. 1980. Practical aspects of stereochemically restrained refinement of protein structures. In *Refinement of protein structures: SERC report* (ed. J.W. Machin et al.), p. 1. Science and Engineering Research Council, Daresbury Laboratory, Warrington, United Kingdom.

Howard, A.J., G.L. Gilliland, B.C. Finzel, T.L. Poulos, D.H. Ohlendorf, and F.R. Salemme. 1987. The use of an imaging proportional

counter in macromolecular crystallography. *J. Appl. Crystallogr.* **20:** 383.
James, M.N.G. and A. Sielecki. 1983. Structure and refinement of penicillopepsin at 1.8 Å resolution. *J. Mol. Biol.* **163:** 299.
Jones, T.A. 1978. A graphics model building and refinement system for macromolecules. *J. Appl. Crystallogr.* **11:** 268.
Katoh, I., Y. Yoshiyuki, A. Rein, M. Shibuya, T. Odaka, and S. Oroszlan. 1985. Murine leukemia virus maturation: Protease region required for conversion from "immature" to "mature" core form and for virus infectivity. *Virology* **145:** 280.
Miller, M., M. Jaskólski, J.K.M. Rao, J. Leis, and A. Wlodawer. 1989. Crystal structure of a retroviral protease proves relationship to aspartic protease family. *Nature* **337:** 576.
Moelling, K.A., K.E. Scott, K.E. Dittmar, and M. Owada. 1980. Effect of P-15 associated protease from an avian RNA tumor virus on avian virus-specific polyprotein precursors. *J. Virol.* **33:** 680.
Navia, M.A., P.M.D., Fitzgerald, B.M. McKeever, C.-T. Leu, J.C. Heimbach, W.K. Herber, I.S. Sigal, P.L. Darke, and J.P. Springer. 1989. Three-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1. *Nature* **337:** 615.
Pearl, L.H. and W.R. Taylor. 1987. A structural model for the retroviral proteases. *Nature* **329:** 351.
Ratner, L., W. Haseltine, R. Patarca, K.J. Livak, B. Starcich, S.F. Josephs, E.R. Doran, J.A. Rafalski, E.A. Whitebourne, K. Baumeister, L. Ivanoff, S.R. Petteway, Jr., M.L. Pearson, J.A. Lautenberger, T.S. Papas, J. Ghrayeb, N.T. Chang, R.C. Gallo, and F. Wong-Staal. 1985. Complete nucleotide sequence of the AIDS virus, HTLV-III. *Nature* **313:** 277.
Sedlacek, J., P. Strop, F. Kapralek, V. Pecenka, V. Kostka, M. Trannicek, and J. Riman. 1988. Processed enzymatically active protease (P 15 *gag*) of avian retrovirus obtained in an *E. coli* system expressing a recombinant precursor (Pr25 lac-Δgag). *FEBS Lett.* **237:** 187.
Toh, H., M. Ono, K. Saigo, and T. Miyata. 1985. Retroviral protease-like sequence in the year transposon Ty1. *Nature* **315:** 691.
Weber, P.C., D.H. Ohlendorf, J.J. Wendoloski, and F.R. Salemme. 1989. Structural origins of high-affinity Biotin binding to Streptavadin. *Science* **243:** 85.

Structural Comparisons of Retroviral and Eukaryotic Aspartic Proteinases

J. Erickson,[1] J.K.M. Rao,[2] C. Abad-Zapatero,[1] and A. Wlodawer[2]

[1]Protein Crystallography Laboratory, D-47E
Abbott Laboratories
Abbott Park, Illinois 60064

[2]Crystallography Laboratory, NCI-Frederick Cancer
Research Facility, BRI-Basic Research Program
Frederick, Maryland 21701

Retroviruses, such as the human immunodeficiency virus (HIV), encode a proteinase that is responsible for the maturation of virions into infectious virus particles (Crawford and Goff 1985; Katoh et al. 1985; Kohl et al. 1988). Maturation proceeds via the specific cleavage of the viral polyprotein products of the *gag* and *pol* genes by the viral proteinase. Much effort is currently being directed toward the design of specific inhibitors of the HIV proteinase enzyme (HIV PR) with the hope that this strategy will lead to a chemotherapeutic agent for the treatment of AIDS. Recent X-ray crystallographic investigations of the Rous sarcoma virus proteinase (RSV PR) (Miller et al. 1989) and HIV PR (Navia et al. 1989) have demonstrated that these enzymes are structurally related to the aspartic proteinase family, for which at least 22 amino acid sequences are known (Foltmann 1988). In addition, X-ray crystallographic structures of three mammalian aspartic proteinases, pepsin (Andreeva et al. 1984; C. Abad-Zapatero et al., in prep.), chymosin (G. Gilliland et al., in prep.), and human renin (Sielecki et al. 1989); and of three fungal enzymes, penicillopepsin (James and Sielecki 1983), rhizopuspepsin (Suguna et al. 1987), and endothiapepsin (Blundell et al. 1985); have been solved and refined at or below 2.5 Å resolution.

Whereas the mammalian and fungal enzymes all possess very similar three-dimensional structures (Tang et al. 1978; C. Abad-Zapatero et al., in prep.), the retroviral proteinases show only limited homology with their eukaryotic counterparts. For example, the eukaryotic enzymes consist of about 325 residues

Current Communications
IN MOLECULAR BIOLOGY
CSH Cold Spring Harbor Laboratory Press

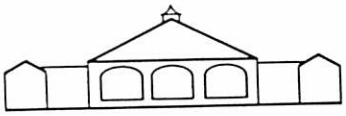
Viral Proteinases as Targets for Chemotherapy

Edited by

Hans-Georg Kräusslich
State University of New York, Stony Brook

Stephen Oroszlan
NCI-Frederick Cancer Research Facility

Eckard Wimmer
State University of New York, Stony Brook



A Banbury Center Meeting