dum-normative spinel cumulate.

In conclusion, we postulate that metamorphic recrystallization of a spinel-rich igneous cumulate at a maximum pressure of 2.5 kbar and depth of 50 km in the lunar crust produced the cordierite-spinel troctolite. As the crust-mantle boundary is at a depth of 60 km and the magnesium-rich suite probably is the most abundant crustal material at depths greater than about 30 km, the rock must have originated at a depth of 30 to 50 km. At such levels, spinel cumulates may well be more important components of the lunar crust than has been recognized. The rarity of cordierite-spinel troctolite indicates that it forms from spinel-rich precursors at horizons sufficiently deep to be unroofed only in major impact events. We suggest that the clast of cordierite-spinel troctolite was excavated from the adjacent Imbrium Basin.

REFERENCES AND NOTES

1. M. Prinz, E. Dowty, K. Keil, T. E. Bunch, *Science* **179**, 74 (1973).
2. A. L. Albee *et al.*, *Lunar Sci.* **5** (Part 1), 3 (1974).
3. A. T. Anderson, *J. Geol.* **81**, 219 (1973).
4. A. E. Bence, J. W. Delano, J. J. Papike, K. L. Camperon, *Proceedings of the Fifth Lunar Science Conference*, J. Dorman *et al.*, Eds. (Pergamon, New York, 1974), pp. 785–827.
5. R. F. Dymek, A. L. Albee, A. A. Chodos, *Proceedings of the Seventh Lunar Science Conference*, R. B. Merrill *et al.* Eds. (Pergamon, New York, 1976), pp. 2335–2378.
6. C. T. Herzberg, *Contrib. Mineral. Petrol.* **84**, 844 (1983).
7. ——— and M. B. Baker, in *Proceedings of the Conference on the Lunar Highlands Crust*, R. B. Merrill, Ed. (Pergamon, New York, 1980), p. 113.
8. S. Arai, *Contrib. Mineral. Petrol.* **52**, 1 (1975).
9. L. H. Fuchs, *Am. Mineral.* **54**, 1645 (1969).
10. G. A. Chinner and J. F. Schairer, *Am. J. Sci.* **260**, 611 (1962).
11. C. D. Gribble and M. J. O'Hara, *Nature* **214**, 1201 (1967).
12. D. C. Presnall *et al.*, *Contrib. Mineral. Petrol.* **66**, 203 (1978).
13. P. H. Warren *et al.*, *J. Geophys. Res.* **88**, B151 (1983).
14. M. L. Lindstrom, U. B. Marvin, D. W. Mittelfeldt, *Proc. 19th Lunar Planet. Sci. Conf.*, in press.

15. We thank E. Roedder and C. Herzberg for helpful critiques of the manuscript. This research was supported in part by NASA grant NAG 9-29 to U.B.M., DOE grant DE-FG02-87-ER13799 to J. B. Thompson, Jr., of Harvard University, and NASA support of M.M.L.

6 October 1988; accepted 21 December 1988

Molecular Modeling of the HIV-1 Protease and Its Substrate Binding Site

IRENE T. WEBER, MARIA MILLER, MARIUSZ JASKÓLSKI,* JONATHAN LEIS, ANNA MARIE SKALKA, ALEXANDER WLODAWER

The human immunodeficiency virus (HIV-1) encodes a protease that is essential for viral replication and is a member of the aspartic protease family. The recently determined three-dimensional structure of the related protease from Rous sarcoma virus has been used to model the smaller HIV-1 dimer. The active site has been analyzed by comparison to the structure of the aspartic protease, rhizopuspepsin, complexed with a peptide inhibitor. The HIV-1 protease is predicted to interact with seven residues of the protein substrate. This information can be used to design protease inhibitors and possible antivirual drugs.

THE LIFE CYCLE OF RETROVIRUSES such as HIV, the pathogen of AIDS, requires a specific protease (PR) that processes the precursor *gag* and *pol* polyproteins into mature virion components (*1–3*). If the PR is absent or inactive, noninfectious virus particles with aberrant structure are produced (*4*). Therefore, specific inhibitors of retroviral proteases are potential therapeutic agents for blocking HIV infection. However, rational drug design requires detailed knowledge of the three-dimensional structure of the target. Recently we solved the crystal structure of the PR from Rous sarcoma virus (RSV), Prague-C strain (*5*). We have used this structure to construct a model of HIV-1 PR and present details of the proposed substrate-binding site of the HIV-1 enzyme.

All of the retroviral proteases contain a triplet, Asp-Thr/Ser-Gly, which, together with known biochemical properties of these enzymes, indicates that they are members of the aspartic protease family (*6–8*). This relation was confirmed by the three-dimensional structure of the RSV PR (*5*), which has now been refined to an *R*-factor of 0.16 with 2 Å x-ray data. Two PR subunits interact to form a symmetric dimer with an active site that is similar to the highly conserved active sites of monomeric cellular proteases such as pepsin (*9*), rhizopuspepsin (*10*), and endothiapepsin (*11*). The root-mean-square (rms) deviation in the positions of 36 superimposed Cα atoms near the active site is about 2 Å between rhizopuspepsin (*10*) and the RSV PR dimer (*5*). There are also considerable similarities in the overall shape of the two molecules. For 170 Cα atoms aligned in both structures (out of 232 placed in the current RSV PR model), the total rms deviation is only 3.2 Å (*12*). Retroviral proteases, however, are simpler than the cell-derived pepsin-like proteins. The dimer of the retroviral PR consists of 200 to 250 amino acids, whereas a single molecule of a cellular protease is about 325 amino acids long and exhibits intramolecular quasi-symmetry. The structural elements of cellular aspartic proteases are conserved

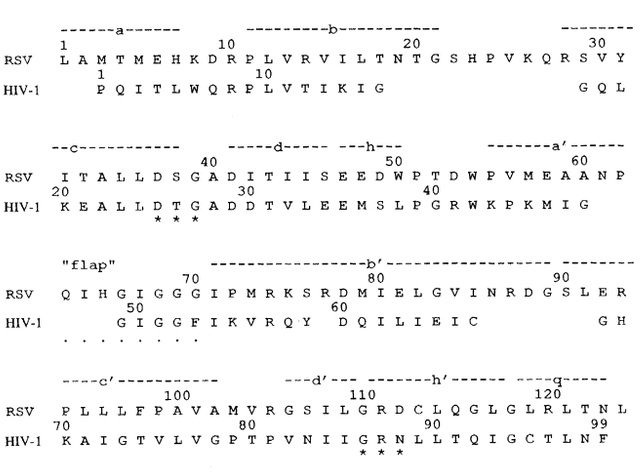

Fig. 1. Amino acid sequence alignment for the proteases from HIV-1 and RSV (*22*). The secondary structural elements of the RSV PR are indicated on top of the amino acid sequences and are labeled according to the scheme proposed by Blundell *et al.* (*11*). A one-turn α helix is labeled *h* helix, and a longer α helix is labeled *h'*. The flexible flap lies between β strands *a'* and *h'*, and residues 63 to 70 were not visible in the electron density map for RSV PR (dotted line). This alignment conserves the topology, so that deletions are positioned in surface loops of the three-dimensional structure. The active-site triad Asp-Thr/Ser-Gly is indicated by asterisks, as is the conserved triplet Gly-Arg-Asp/Asn at the start of helix *h'*. The flap region was aligned based on the rules proposed by Andreeva (*23*).

I. T. Weber, M. Miller, M. Jaskólski, A. Wlodawer, Crystallography Laboratory, NCI-Frederick Cancer Research Facility, BRI-Basic Research Program, Frederick, MD 21701.
J. Leis, Department of Biochemistry, Case Western Reserve University, Cleveland, OH 44106.
A. M. Skalka, Fox Chase Cancer Center, Institute for Cancer Research, Philadelphia, PA 19111.

*On leave from A. Mickiewicz University, Faculty of Chemistry, Poznan, Poland.

and this feature allowed modeling of other members of the family (*13*).

A model of the HIV-1 PR has been built by analogy to the RSV PR structure. The active site triplet and a second conserved sequence, Gly-Arg-Asn/Asp, which lies toward the carboxyl terminus at the start of the α helix in the RSV PR structure, were used to align the sequences of the two proteins (Fig. 1). This alignment results in 30 identical residues and 11 additional conservative amino acid substitutions between the two enzymes. Since the HIV-1 protease has only 99 residues, compared with 124 for the RSV PR, a structural alignment was built into the model by permitting deletions only at the large surface loops. These deletions could be confined to three regions at surface turns between β strands in the RSV PR structure (Figs. 1 and 2). Analogous deletions occur at the same surface turns when the structure of rhizopuspepsin is compared with RSV PR. The HIV-1 model structure also places amino acids commonly found in turns at the appropriate positions. All other proteases from mammalian viruses can be accomodated by a similar alignment, including the longer proteases, such as Mo-MuLV, which has extra residues at its carboxyl terminus. The presence of longer loops in the RSV PR may serve a structural role in virion core assembly since, unlike HIV-1, the RSV PR is synthesized as part of both *gag* and *gag-pol* precursors and is thus present in the virion in concentrations equal to those of the capsid structural proteins.

The structure of the modeled HIV-1 PR dimer is illustrated in Fig. 3. The overall topology of this and of the original RSV PR structure is similar to the model proposed by Pearl and Taylor (*14*). Their model was constructed by analogy only to cellular aspartic proteases, and the exact positions of the elements of secondary and tertiary structure differ from that in the RSV PR structure. The model we propose should be more accurate since it is based upon the more closely related retroviral (RSV PR) structure. The specific cleavage by RSV PR of some HIV-1 peptide substrates (*15*) supports this rationale. The active site cleft of the HIV-1 PR predicted in this model includes amino acid residues 23 to 32, which contain the active site triplet, residues 49 to 59 in the "flap" region, residues 74 to 76 and 80 to 87, which contain part of a β sheet, and the α helix region near the carboxyl terminus. These regions of the polypeptide chain would be predicted to be the most sensitive to inactivation by mutations since changes in the component amino acids might easily disrupt the active site of the PR. Debouck *et al.* (*2*) introduced a 4–amino-acid insertion of Ile^{54} in the flap region of the HIV-1 PR, which inactivated the enzyme. Loeb *et al.* (*16*) introduced random mutations into selected regions of the HIV-1 PR and found that amino acid residues 22 to 28 were very sensitive to substitutions, whereas residues 34 to 46 were not. Finally, all of the amino acid substitutions made at the catalytic Asp^{25} have resulted in an inactive PR (*3, 7, 16, 17*). The model proposed here is consistent with all of these results.

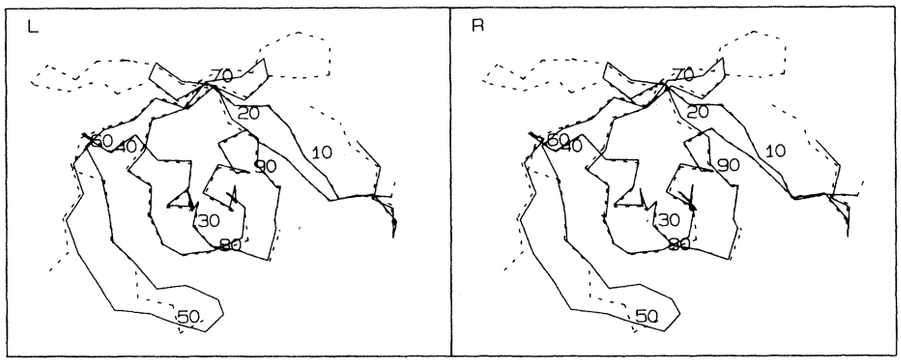

Fig. 2. A stereoview of the Cα chain of the model HIV-1 PR (continuous lines) superimposed on the Cα chain of the crystallographically determined structure of the RSV protease (*5*) (dashed lines). Every tenth residue of the HIV-1 PR is labeled. The extended β turns, *b* to *c*, and *b'* to *c'*, are shortened in the HIV-1 model relative to the RSV protease structure. The flap is also five residues shorter in HIV-1 PR. The atomic coordinates of the model of HIV-1 PR described here are available from the authors.

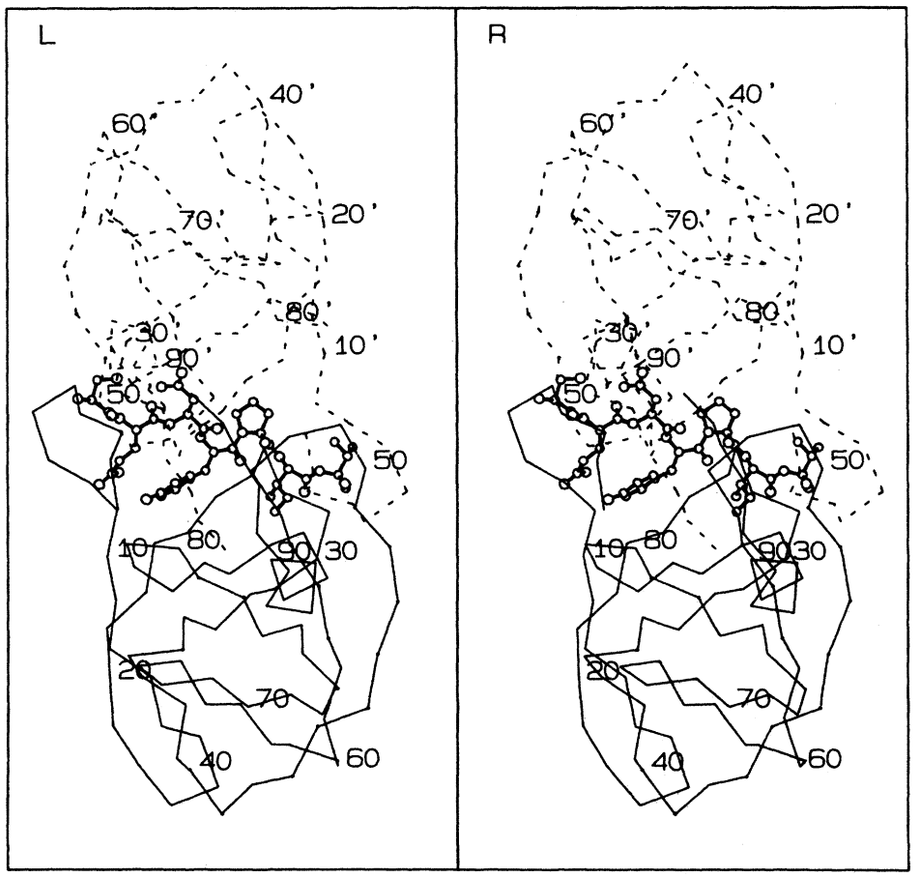

Fig. 3. A stereoview of the active dimer of the model HIV-1 PR with the substrate, Ser-Gln-Asn-Tyr↓Pro-Ile-Val. The Cα chains of the two subunits in the dimer are indicated by continuous and dashed lines and every tenth residue is numbered. The terminology for aspartic protease binding sites has been adopted for discussion of the modeled retroviral PR substrate binding sites. Substrate residues P4 to P1 before the scissile bond and P1′ to P3′ after the scissile bond are shown. Substrate side chains bind in successive protease subsites, so residue P1 lies in subsite S1. The seven-residue substrate (P4 to P3′) is positioned with P4 to P1′ as in the rhizopuspepsin-inhibitor complex, residue P2′ is moved slightly and residue P3′ added. The substrate is indicated by dark lines in a ball-and-stick representation.

The rhizopuspepsin structure complexed with the reduced peptide inhibitor, His-Pro-Phe-His-Phe [CH–NH]Phe-Val-Tyr (where the reduced bond is between the two Phe residues) (*10*), has been superimposed on the HIV-1 PR model structure. The active sites of these enzymes closely resemble one another. Both pepstatin analogs and the reduced octapeptide inhibitor bind to the cellular proteases in an extended β conformation (*10, 18*). The side chains of the peptide inhibitors lie in successive subsites in the protease binding cleft. In rhizopuspepsin, six residues of the reduced octapeptide inhibitor are observed within subsites S1 and S4 before the scissile peptide bond, and subsites S1′ and S2′ after the scissile bond. The two catalytic Asp residues lie near the scissile bond, and several hydrogen bond interactions occur between the protease and the amides and carbonyl oxygens of the inhibitor peptide (*10*).

Kotler *et al.* (*15*) found that the minimal size of the RT α-IN protein junction peptide substrate for the RSV PR is between three and five amino acids to either side of the scissile bond. Darke *et al.* (*19*), using peptides corresponding to a target in the HIV-1 *gag* precursor, have shown that a hepta- but not hexapeptide is cleaved by the HIV-1 PR. Thus it is likely that a minimal target sequence for retroviral proteases would include four amino acids (P4–P1) corresponding to the side that is to become the new carboxyl terminus, and three amino acids (P1′–P3′) corresponding to that which is to become the new amino terminus of the cleaved product.

An HIV-1 peptide substrate, Ser-Gln-Asn-Tyr↓Pro-Ile-Val, which is cleaved efficiently (*20*) (at the bond shown by the arrow), was thus modeled by analogy to the reduced octapeptide inhibitor of rhizopuspepsin in order to identify, as far as possible, the protein subsites for each substrate residue (Fig. 3). The HIV-1 PR catalytic residues, Asp^{25} and $Asp^{25'}$ from both subunits in the dimer, are near the scissile bond. Potential hydrogen bond interactions between the HIV-1 PR and the substrate amides and carbonyl oxygens are listed in Table 1. Similar interactions are formed in the model of HIV-1 PR with substrate compared with those observed in the crystal structures of rhizopuspepsin complexed with inhibitor (*10*) and of penicillopepsin with a pepstatin analog inhibitor (*18*). This similarity suggests that the conformation of the peptide substrate in the model with HIV-1 PR is correct.

The exact position of the two flexible flaps in the active dimer of PR is not known, and may very well be altered in the presence of the substrate. The structures of rhizopuspepsin and penicillopepsin complexed with inhibitors (*10, 18*) show a major conformational change of the flap compared with the uncomplexed enzymes. This change enhances binding of the inhibitor by facilitating several hydrogen bond interactions; it also causes an exclusion of water from the catalytic site. It is assumed that the flap of retroviral proteases has a similar function, although it is not known whether one or both flaps in the dimer are involved. It is possible that the presence of two flexible flaps allows binding of the variety of sequences in the polyprotein precursors known to be cleaved by the retroviral PR.

Seven protease subsites were examined (Table 2) in order to identify differences between HIV-1 and RSV proteases and rhizopuspepsin. The proposed retroviral PR subsites have been correlated with known HIV-1 protease substrate cleavage sites. Since the HIV-1 and RSV PR target sequences are variable (*19, 21*) it is difficult to define a consensus sequence for cleavage. However, substrate residues P1 and P1′, which are on either side of the scissile bond, are almost always hydrophobic. The HIV-1 PR is most active against the peptide Ser-Phe-Asn-Phe↓Pro-Gln-Ile-Thr (*19*), which corresponds to the site at which cleavage releases the amino terminus of the PR from the precursor polypeptide. A smaller residue would be predicted to bind in subsite S4. There is the potential to bind polar residues at substrate residue P3 since both retroviral enzymes have the conserved Asp and Arg at positions corresponding to residues 29 and 8 of the HIV-1 PR. However, the HIV-1 PR substrate with the fastest rate of hydrolysis, as noted above, contains Phe at P3. The

Table 1. Potential hydrogen bond interactions between HIV-1 PR or rhizopuspepsin (*10*) and the amides and carbonyl oxygens of their peptide substrates. Residues from the two subunits in the dimer of HIV-1 PR are distinguished by a prime, for example, Asp^{25} and $Asp^{25'}$. The scissile bond is cleaved between P1 and P1′.

Subsite	Substrate atom	HIV-1 PR atom	Rhizopuspepsin atom
S4	P4 NH		
†	P4 C=O	NH $Gly^{51'}$	
S3	P3 NH	O $Asp^{29'}$	OH Thr^{222}
	P3 C=O	NH $Asp^{29'}$	NH Thr^{222}
S2†	P2 NH	C=O $Gly^{51'}$	O Asp^{79}
†	P2 C=O	NH $Phe^{53'}$	NH Gly^{78}
S1*	P1 NH	C=O $Gly^{27'}$	C=O Gly^{220}
*	P1 C=O	O Asp^{25}	O Asp^{218}
S1′*	P1′ NH	O $Asp^{25'}$	O Asp^{35}
†	P1′ C=O	NH Phe^{53}	NH Gly^{78}
S2′*	P2′ NH	C=O Gly^{27}	C=O Gly^{37}
	P2′ C=O	NH Asp^{29}	NE Trp^{194}
S3′†	P3′ NH	C=O Gly^{51}	
†	P3′ C=O	NH Gly^{51}	

*Residues from the active site triad, Asp-Thr-Gly, make these interactions. †Residues from the flap form these interactions, although the exact conformation of the flaps in HIV-1 PR is difficult to predict.

Table 2. Protease residues near the amino acid side chains of the substrate peptide.

Subsite	HIV-1 PR	Protease residues RSV PR(*5*)	Rhizopuspepsin (*10*)
S4	$Asp^{30'}$ $Ile^{50'}$ $Ile^{54'}$	$Ile^{42'}$ $Ile^{71'}$ $Met^{73'}$	Thr^{222} Leu^{223} Ile^{15}
S3	Arg^{8} $Asp^{29'}$ Leu^{23} Val^{82} $Arg^{87'}$	Arg^{10} $Asp^{41'}$ Arg^{105} Leu^{35} $Arg^{111'}$	Tyr^{11} Ile^{15} Glu^{16} Pro^{118} Asn^{119}
S2	$Ala^{28'}$ Val^{32} $Phe^{53'}$ $Ile^{54'}$ $Leu^{76'}$ $Thr^{80'}$ $Ile^{84'}$	$Ala^{40'}$ Ile^{44} $Met^{73'}$ $Val^{104'}$ $Ile^{108'}$	Gly^{78} Asp^{79} Thr^{221} Ile^{225}
S1	Leu^{23} Asp^{25}* Phe^{53} Pro^{81} Val^{82} Ile^{84}	Leu^{35} Asp^{37}* Val^{104} Arg^{105} Ile^{108}	Asp^{33} Asp^{35}* Tyr^{77} Asp^{79} Phe^{114} Leu^{122} Asp^{218}*
S1′	$Leu^{23'}$ Asp^{25}*′ $Phe^{53'}$ $Ile^{84'}$	$Leu^{35'}$ Asp^{37}*′ $Val^{104'}$ $Arg^{105'}$ $Ile^{108'}$	Trp^{194} Ile^{216} Asp^{218}* Trp^{294} Ile^{298}
S2′	Ala^{28} Asp^{30} Val^{32} Phe^{53} Ile^{54} Leu^{76} Ile^{84}	Ala^{40} Ile^{42} Ile^{44} Val^{104} Ile^{108}	Ser^{38} Ile^{75} Tyr^{77} Ile^{130}
S3′	$Arg^{8'}$ Asp^{29} Ile^{50} $Pro^{81'}$ Arg^{87}	$Arg^{10'}$ Asp^{41} $Arg^{105'}$ Arg^{111}	

*Catalytic Asp residue.

extended β conformation of the peptide substrate would permit favorable interactions between Phe at both P3 and P1 as may occur in the rhizopuspepsin inhibitor complex.

Subsite S1 in these proteases has space for a large hydrophobic residue, which agrees with experimental observations with RSV PR showing that iodinated Tyr at P1 is cleaved more rapidly than Tyr or smaller hydrophobic residues (*8*). The Phe residue occurs at P1 in several HIV-1 cleavage sites, and there may be favorable interactions with Phe^{53}. In rhizopuspepsin, Phe may occur at P1 and P1′ due to favorable interactions with Tyr^{77} and Phe^{114} of subsite S1 and Trp^{194} and Trp^{294} in S1′. Although rhizopuspepsin binds Phe at P1′, a smaller hydrophobic side chain is predicted to fit better in subsite S1′ of the retroviral PR and Pro is often present at this location in both RSV and HIV-1 target sequences.

The substrate residue P2′ tends to be polar for HIV-1 PR and hydrophobic for RSV PR. This difference may be due to the nearby Asp^{30} in HIV-1 PR, which is replaced by a hydrophobic residue Ile^{42} in RSV PR. Further from the cleaved peptide, that is, at and beyond subsites S4 and S3′, the substrate lies near the surface of the dimer where there are several charged residues, so polar amino acids may be preferred at distal locations in the substrate. The above subsite model and other predictions from our analyses are presently being tested by site-directed mutagenesis of the PR and substrate peptides. However, these predictions provide sufficient insight into the probable features of the active site of the HIV-1 protease to be immediately useful for the design of potential inhibitors of this retroviral enzyme.

REFERENCES AND NOTES

1. H.-G. Kräusslich and E. Wimmer, *Annu. Rev. Biochem.* **57**, 701 (1988); W. C. Farmerie *et al.*, *Science* **236**, 305 (1987); R. A. Kramer, *ibid.* **231**, 1580 (1986); M. C. Graves, J. J. Lim, E. P. Heimer, R. A. Kramer, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 2449 (1988).
2. C. Debouck *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 8903 (1987).
3. J. Mous, E. P. Heimer, S. F. J. LeGrice, *J. Virol.* **62**, 1433 (1988).
4. I. Katoh *et al.*, *Virology* **145**, 280 (1985); S. Crawford and S. Goff, *J. Virol.* **53**, 899 (1985).
5. M. Miller, M. Jaskolski, J. K. M. Rao, J. Leis, A. Wlodawer, *Nature*, in press.
6. I. Katoh, T. Yasunaga, Y. Ikawa, Y. Yoshinaka, *ibid.* **329**, 654 (1987); H. Toh, M. Ono, K. Saigo, T. Miyata, *ibid.* **315**, 691 (1985); R. H. Miller, *Science* **236**, 722 (1987); J. Schneider and S. Kent, *Cell* **54**, 363 (1988); R. F. Nutt *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 7129 (1988); J. Hansen, S. Billich, T. Schulze, S. Sukrow, K. Moelling, *EMBO J.* **7**, 1785 (1988); H.-G. Kräusslich, H. Schneider, G. Zybarth, C. Carter, E. Wimmer, *J. Virol.* **62**, 4393 (1988); J. Tang, M. N. G. James, I. N. Hsu, J. A. Jenkins, T. L. Blundell, *Nature* **271**, 618 (1978).
7. S. Seelmeier, H. Schmidt, V. Turk, K. von der Helm, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6612 (1988).
8. M. Kotler, W. Danho, R. A. Katz, J. Leis, A. M. Skalka, *J. Biol. Chem.*, in press.
9. N. S. Andreeva, A. J. Zdanov, A. E. Gustchina, A. A. Fedorov, *ibid.*, **259**, 11353 (1984).
10. K. Suguna *et al.*, *J. Mol. Biol.* **196**, 877 (1987).
11. T. Blundell, J. Jenkins, L. Pearl, T. Sewell, V. Pedersen, in *Aspartic Proteinases and Their Inhibitors*, V. Kostka, Ed. (de Gruyter, Berlin, 1985), pp. 151–161.
12. I. T. Weber *et al.*, unpublished observations.
13. T. Blundell, B. L. Sibanda, L. Pearl, *Nature* **304**, 273 (1983).
14. L. Pearl and W. Taylor, *ibid.* **329**, 351 (1987).
15. M. Kotler, R. Katz, W. Danho, J. Leis, A. Skalka, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 4185 (1988).
16. D. D. Loeb, C. A. Hutchison, M. H. Edgell, W. G. Farmerie, R. Swanson, *J. Virol.* **63**, 111 (1989).
17. M. Kotler *et al.*, *ibid.* **62**, 2696 (1988).
18. M. N. G. James, A. R. Sielacki, J. Moult, in *Peptides: Structure and Functions, Proceedings of the 8th American Peptide Symposium*, V. Hornby and D. Rich, Eds. (Pierce Chemical, Rockford, IL, 1983), pp. 521–530.
19. P. L. Darke *et al.*, *Biochem. Biophys. Res. Commun.* **156**, 297 (1988).
20. S. Billich *et al.*, *J. Biol. Chem.* **263**, 17905 (1988).
21. L. E. Henderson *et al.*, *J. Virol.* **62**, 2587 (1988).
22. Abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
23. N. S. Andreeva, in *Structure and Biosynthesis of Proteins* (in Russian) (Soviet Academy of Sciences, Pushchino, 1988), vol. 3, pp. 3–21.

24. We thank N. S. Andreeva for invaluable discussions and R. Swanstrom for providing data before publication. Research sponsored by the National Cancer Institutes, DHHS, under contract NO1-CO-74101 with Bionetics Research, Inc. J.L. is supported in part by Public Health Service grant CA38046 from the National Cancer Institute and in part by a pilot research grant from the American Cancer Society Cuyahoga County Division. A.M.S. is supported by National Institutes of Health grants CA-06927 and RR-05539, a grant from the Pew Charitable Trust, and also by an appropriation from the Commonwealth of Pennsylvania. The contents of this publication do not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. government.

12 December 1988; accepted 27 January 1989

Ubiquitous Expression of *sevenless*: Position-Dependent Specification of Cell Fate

KONRAD BASLER AND ERNST HAFEN

Specification of cell fate in the compound eye of *Drosophila* appears to be controlled entirely by cell interactions. The *sevenless* gene is required for the correct determination of one of the eight photoreceptor cells (R7) in each ommatidium. It encodes a transmembrane protein with a tyrosine kinase domain and is expressed transiently on a subpopulation of ommatidial precursor cells including the R7 precursors. It is shown here that heat shock–induced indiscriminate expression of a *sevenless* complementary DNA throughout development can correctly specify R7 cell identity without affecting the development of other cells. Furthermore, discontinuous supply of *sevenless* protein during eye development leads to the formation of mosaic eyes containing stripes of *sevenless*$^+$ and *sevenless*$^-$ ommatidia, suggesting that R7 cell fate can be specified only within a relatively short period during ommatidial assembly. These results support the hypothesis that the specification of cell fate by position depends on the interaction of a localized signal with a receptor present on many undifferentiated cells, and that the mere presence of the receptor alone is not sufficient to specify cell fate.

THE COMPOUND EYE OF *DROSOPHILA* is a suitable model system to study specification of cell fate determined by cellular interactions since the assembly of the individual ommatidia or unit eyes occurs without cell lineage restrictions (*1–3*). It has been proposed that the fate of an undetermined cell depends on its position relative to previously determined cells in the developing eye (*4*). The *sevenless* (*sev*) gene is required for the correct determination of one of the eight photoreceptor cells (R7) in each ommatidium (*5*). In the absence of a functional *sev* gene the presumptive R7 cell does not enter the proper developmental pathway and assumes a different fate (*6*). The nature of the *sev* gene product, which appears to be a receptor-type tyrosine kinase, suggests that it interacts directly with a ligand presented by a neighboring cell. This interaction would then lead to the determination of the cell through the activation of the *sev* protein (*7, 8*). Specification of cell fate by cell interactions can be viewed in at least two alternative ways. Specificity of the selection may depend on the localized presentation of a positional signal and a widespread distribution of the corresponding receptor. Conversely, selection might be dictated by the restricted distribution of the receptor interacting with a ubiquitous ligand. Although the *sev* protein is not exclusively expressed in the R7 precursor, its expression is temporally and spatially restricted within the eye disk (*9*). Therefore it

Zoologisches Institut, Universität Zürich, CH-8057 Zürich, Switzerland.