C2/C3–C6′ in Fig. 2 and vector Se1–Se3″ and Se2″–Se4 in Fig. 3). Figure 2 is a "slice" taken through the bisector of the TMTSF stack (mirror plane bisecting bonds C7–C8 and C2–C3 along the C1–C6 bond), and Fig. 3 is a slice through a plane made up of selenium atoms on the right-hand side of a stack of TMTSF molecules (see Fig. 1).

Examination of Figs. 2 and 3 shows considerable electron density in an alternating pattern between TMTSF molecular centers and between selenium atoms at the edge of adjacent stacks. The alternating intrastack density appears to be an extension of the possible transannular selenium bonding or overlap of appropriate Se-C-Se atomic orbitals and could be considered a microscopic view of electron density distribution along a one-dimensional conduction band.

In order to locate electron density between stacks, it was assumed that electron transfer would be accomplished through the selenium atoms, as suggested earlier (*8*). Therefore, electron densities in the interstack planes a and c of Fig. 5 were calculated and are depicted in Fig. 4, a and b, respectively.

Surprisingly, and contrary to interactions depicted previously (*8*), electron density was observed only along the longest Se-Se interstack distance (*15*) (Se2–Se4′, 4.151 Å) as shown in Fig. 4a defined by the atoms Se1′, Se4′, Se2, and Se3. Aside from the repulsive force of the opposing Se-Se lone-pair density, there is a continuum of density from one molecule to the other which may represent a conduction band. Similar interstack bonding through sulfur lone pairs was postulated for $(SN)_x$ (*16*). In contrast, in the shortest selenium atom contact (Se2–Se2′, 3.905 Å), no electron density is observed in the intermolecular plane (*11*). The next largest interstack selenium distance is 3.963 Å and corresponds to Se4–Se4′. In this case, even though the selenium atom lone pairs are pointed toward each other, there is no intermolecular continuum of density (Fig. 4b, plane defined by atoms Se1, Se4 and Se1′, Se4′).

We have shown that there exists intermolecular bonding both within stacks and between stacks in $(TMTSF)_2AsF_6$ and $(TMTSF)_2PF_6$. The intermolecular bonding electrons originate from the selenium atoms. One of the proposed four possible modes of interstack selenium bonding is supported by this experimental determination of electron density represented schematically in Fig. 5 (*17*).

Our results are in accord with previous suggestions concerning the dimensionality of the Bechgaard phase and observations of magnetotransport (*18*) in $(TMTSF)_2PF_6$. This experimental electron density study of an organic metal shows bonding and lone-pair density for selenium atoms along with possible conduction band electrons.

Table 1. Intermolecular Se–Se contacts to 5.0 Å.

Atom 1	Atom 2	Distance	Symmetry code*
Se1	Se2	4.053	(−1 1 1 1)
Se1	Se2	4.128	(−1 2 1 1)
Se1	Se3	4.026	(−1 1 1 1)
Se1	Se3	3.874	(−1 2 1 1)
Se1	Se4	3.945	(−1 1 2 1)
Se2	Se2	3.905	(−1 2 0 1)
Se2	Se3	4.977	(−1 2 1 1)
Se2	Se4	4.151	(−1 0-1 0)
Se2	Se4	3.970	(−1 1 1 1)
Se2	Se4	3.919	(−1 2 1 1)
Se4	Se4	3.963	(−1 1 2 1)

*The symmetry code *n i j k* denotes the symmetry operation *n** (*x, y, z*) + (*i, j, k*). For example, the code (−1 1 1 1) corresponds to (−*x*, −*y*, −*z*) + (1, 1, 1) or (1 − *x*, 1 − *y*, 1 − *z*). Intrastack close contacts will have codes of (−1 1 1 1) or (−1 2 1 1). The remaining codes are for various interstack contacts.

F. WUDL*
D. NALEWAJEK†
Bell Laboratories,
Murray Hill, New Jersey 07974
J. M. TROUP
M. W. EXTINE
Molecular Structure Corporation,
3304 Longmire Drive,
College Station, Texas 77840

References and Notes

1. K. Bechgaard and D. Jerome, *Sci. Am.* **247**, 52 (July 1982).
2. R. M. Metzger, *Mol. Cryst. Liq. Cryst.* **85**, 57 (1982).
3. A. N. Bloch, cited in (*2*); S. Mazumdar and A. N. Bloch, preprint.
4. P. Coppens, *Int. Rev. Sci.; Inorg. Chem. Ser. Two* **11** 21 (1975).
5. J. M. Troup and R. F. Ziolo, *J. Am. Chem. Soc.* **105**, 229 (1983).
6. M. W. Extine and J. M. Troup, in preparation.
7. J. M. Troup, M. W. Extine, R. F. Ziolo, in *Proceedings of the American Chemical Society Symposium on Electron Distribution and the Chemical Bond*, M. B. Hall and P. Coppens, Eds. (Plenum, New York, 1982), p. 285.
8. F. Wudl, *J. Am. Chem. Soc.* **103**, 7064, (1981).
9. Electron distribution in the inorganic superconductor V_3Si at room temperature has been determined. This solid, however, is not a molecular crystal and all atoms are covalently bonded to each other. J.-L. Standenmann, P. Coppens, J. Muller, *Solid State Commun.* **19**, 29 (1976).
10. Low-temperature data are better in terms of the possible $\sin\theta/\lambda$ cutoff than room temperature data, but the differences are minor and affect only the quantitative aspects of the electron density. In our study, the generation of large peaks around the selenium atoms at low temperature (possibly resulting from thermally induced disorder) is far less desirable than the minor quantitative advantages gained at low temperatures. Some common sense is required in the interpretation of what is an artifact of the experiment and what is "real" electron density. In five crystals examined at different temperatures, some features are always present and are probably correct; other features are created or changed from crystal to crystal at low temperatures and cannot be accepted as real electron density features.
11. Full experimental details including additional electron density maps will be part of a full account to be published elsewhere (F. Wudl, D. Nalewajek, E. Aharon-Shalom, J. M. Troup, M. W. Extine, in preparation).
12. D. W. J. Cruickshank, *Acta Crystallogr.* **2**, 65 (1949).
13. An argument could be made that sp^2 and sp^3-like electron distributions could not be distinguished by examining electron densities because one could take a linear combination of atomic orbitals to give any observed electron density distribution. However, in this case, the absence of electron density in the molecular plane, coupled with the observation of substantial electron density at an angle above and below the plane, is most easily interpretable in terms of symmetry substantially different from sp^2.
14. W. K. Musker, T. L. Wolford, P. B. Roush, *J. Am. Chem. Soc.* **100**, 6416 (1978); K.-D. Asmus, D. Bahnemann, Ch.-H. Fischer, D. Weltwisch, *ibid.* **101**, 5322 (1979). These authors found transannular S–S bonding in radical cations. It has not been established whether *d* orbitals are involved.
15. Computations by P. M. Grant led to the same conclusion, independent of the result reported here [P. M. Grant, *Phys. Rev. B* **26**, 6888 (1982)].
16. M. J. Cohen, A. F. Garito, A. J. Heeger, A. G. Mac Diarmid, C. M. Mikulski, M. S. Saran, J. Klipinger, *J. Am. Chem. Soc.* **98**, 3844 (1976).
17. Figure 2 in (*8*) should be modified accordingly.
18. J. F. Kwak, J. E. Schirber, R. L. Greene, E. M. Engler, *Mol. Cryst. Liq. Cryst.* **79**, 111 (1982).
19. We are particularly indebted to C. K. N. Patel, W. M. Walsh, Jr., and G. A. Thomas for support of this study. We also thank E. A. Chandross for support and Z. Berkovitch-Yellin for critical reading of the manuscript.
* Present address: Department of Physics, University of California, Santa Barbara 93106.
† Present address: Allied Corporation, Buffalo, N.Y. 14210.

28 February 1983; accepted 26 May 1983

Angiotensinogen Is Related to the Antitrypsin-Antithrombin-Ovalbumin Family

Abstract. *The recently reported amino acid sequence of rat angiotensinogen was subjected to a computer-assisted search for homology with known sequences stored in a data bank and found to be significantly related to that of plasma α_1-antitrypsin, itself a member of a family that includes antithrombin III and ovalbumin. An alignment of the four sequences shows indisputably the common ancestry of all four proteins.*

I have been screening all newly published amino acid sequences as they appear in the literature to see if they resemble sequences of proteins already in my collection (*1*). Ohkubo and colleagues (*2*) have recently determined the complementary DNA (cDNA) sequence for the rat angiotensinogen gene and deduced the amino acid sequence of the protein. The gene encodes a preprotein of 477 residues that includes a 24-residue signal peptide. The mature protein, after re-

lease of the signal, has 453 residues. A search of the database revealed a significant resemblance of this protein to human α_1-antitrypsin (A1AT) (*3*, *4*), which itself is known to be related to antithrombin III (AT3H) (*5*) and chicken ovalbumin (*6*, *7*). In all cases the final amino acid sequences have been verified or inferred from cDNA sequences (*3*, *7*, *8*). A comparison of the four sequences leaves no doubt that angiotensinogen is a member of the same family (Fig. 1). The most remarkable aspect of the relationship is that angiotensinogen is the precursor of a very potent biological hormone while A1AT and AT3H are protease inhibitors; the function of ovalbumin is unknown. Angiotensin I is a decapeptide cleaved from the amino terminus of the parent angiotensinogen by the plasma enzyme renin (E.C. 3.4.99.19). The decapeptide is then converted by the angiotensin-converting enzyme (E.C. 3.4.15.1), which removes two amino acids from its carboxyl-terminus, into the more active octapeptide angiotensin II. The angiotensins play an important physiological role in the regulation of blood pressure and water balance (*9*). Angiotensinogen and the active peptides are found both in the circulation and in the central nervous system (*9*).

Table 1. Percent identities (boldface) and statistical significance in standard deviations (italic) of aligned pairs of four sequences (*13*). AGTR, rat angiotensinogen; A1AT, human α_1-antitrypsin; OVCH, chicken ovalbumin; AT3H, human antithrombin III.

Sequence	AGTR	A1AT	OVCH	AT3H
AGTR	—	**23.0**	**20.9**	**17.8**
A1AT	*5.6*	—	**30.8**	**30.8**
OVCH	*5.5*	*16.0*	—	**31.2**
AT3H	*2.5*	*11.6*	*17.9*	—

The case for common ancestry is compelling. The lengths of the four proteins are rat angiotensinogen, 453 residues (*2*); human antithrombin III, 432 residues (*8*); human α_1-antitrypsin, 394 residues (*3*, *4*); and chicken ovalbumin, 385 residues (*7*). As in the case of angiotensinogen, the cDNA sequences of α_1-antitrypsin (*3*) and antithrombin III (*8*) indicate the existence of hydrophobic signal peptides. Ovalbumin, on the other hand, does not have a signal peptide, a single methionine residue being removed from its amino terminus in advance of amino-

```
                                        .........
AGTR      mtptgaglkatifciltwvsltrgDRVYIHPFHLLYYSKSTCAQLENPSVETLPEPTFEPVPIQAKTSPVDEKTLRDKLVLATEKLEAEDRQRAA----
A1AT                                                                    lllaglccllpgslaEDPQGDAAQKTDTSHHDQDHPT-----
OVCH
AT3H                              mysnvigtvtsgkrkvyllslligfwdcvtcHGSPVDICTAKPRDIPMNPMCIYRSPEKKATEDEGSEQKIPEATNRR

                     *       *             *           **    *
AGTR      QVAMIANFMGFR--MYKMLSEARGVASGAVLSPPALFGTLVSFYLGSLDPTASQLQVLLGVPVKEGDCTSRLDGHKVLTALQAVQGLLVTQGGSSSQTPL
A1AT      FNKITPNLAEFAFSLYRQLA-HQSNSTNIFFSPVSIATAFAMLSLGTKADTHDEILEGLNFNLTEIPEAQI---------HEGFQELLRTLNQPDSQL--
OVCH      AcGSIGAASMEFCFDVFKELK-VHHANENIFYCPIAIMSALAMVYLGAKDSTRTQINKVVRFDKLPGFGDSIEAQCGTSVNVHSSLRDILNQITKPNDVY-
AT3H      VWELSKANSRFATTFYQHLADSKNDNDNIFLSPLSISTAFAMTKLGACNDTLQQLMEVFKFDTISEKTSDQIHFFFAKLNCRLYR--------KANKSS-

                *                                                **  *   *                 *          * *     *    *
AGTR      LQSTVVGLFTAPGLRLKQPFVESLG--PFTPAIFPRSLDLSTDPVLAAQKINRFVQAVTGWKMNLPLEGVSTDS-TLFF-NTYVHFQGK-MRGF-SQLTG
A1AT      QLTTGNGLFLSEGLKLVDKFLEDVKKLYHSEAFTVN---FGDTEE-AKKQINDYVEKGTQGKI--VDLVKELDRDTVFALVNYIFFKGKWERPFEVKDTE
OVCH      SFSLASRLYAEERYPILPEYLQCVKELYRGGLEPIN---FQTAADQARELINSWVESQTNGIIRNVLQPSSVDSQTAMVLVNAIVFKGLWEKAFKDEDTQ
AT3H      KLVSANRLFGDKSLTFNETYQDISELVYGAKLQPLD---FKENAEQSRAAINKWVSNKTEGRITDVIPSEAINELTVLVLVNTIYFKGLWKSKFSPENTR

             *           *    * *                          *                                                  *         *
AGTR      LHEFWVDNSTSVSVPMLSGTGNFQHWSDAQNNFSVTRVPLG-ESVTLLLIQPQCASDLDRVEVLVFQHDFLTWIKNPPPRAIR----LTLPQLEIRGSYNL
A1AT      EEDFHVDQVTTVKVPMMKRLGMFNIQHCKKLSSWVL-LMKYLGNATAIFFLPDEGKLQHLENELTHDIITKFLENEDRRSASL-----HLPKLSITGTYDL
OVCH      AMPFRVTEQESKPVQMMYQIGLFRVASMASEKMKILELPFASGTMSMLVLLPDEVSGLEQLES-IINFEKLTEW-TSSNVMEERKIKVYLPRMKMEEKYNL
AT3H      KELFYKADGESCSASMMYQEGKFRYRRVA-EGTQVLELPFKGDDITMVLILPKPEKSLAKVEK-ELTPEVLQEW---LDELEEMMLVVHMPRFRIEDGFSL

             *                                        * *                              **              * *    *
AGTR      QDLLAQAKLSTLLGAEANLGKMGDTN---PRVGE--VLNSILLELQAGEEEQPT---ESAQQPGSPEVLDVTLSSPFLFAIYERDSGALHFLGRVDNPQNVV
A1AT      KSVLGQLGITKVFSNGADLSGVTEEA---PLKLSKAVHKAVLTIDEKGTEAAGA---MFLEAIPMSIPPEVKFNKPFVFLMIEQNTKSPLFMGKVVNPTQK
OVCH      TSVLMAMGITDVFSSSANLSGISSAE---SKLISQAVHAAHAEINEAGREVVGS---AEAGVDAASVSEEFRADHPFLFCIKHIATNAVLFFGRCVSP
AT3H      KEQLQDMGLVDLFSPEKSKLPGIVAEGRDDLYVSDAFHKAFLEVNEEGSEAAASTAVVIAGRSLNPNRVTFKANRPFLVFIREVPLNTIIFMGRVANPCVK
```

Fig. 1. Alignment of four related sequences: AGTR, angiotensinogen; A1AT, α_1-antitrypsin; OVCH, chicken ovalbumin; AT3H, human antithrombin. Asterisk (*) denotes residues in which all four sequences have the same amino acid. Ac, acetylated amino terminal. Dottled line, angiotensin I region. Signal peptide portions are shown in lowercase. Single letter abbreviations for the amino acids are A, alanine; C, cysteine; D, aspartic acid; E, glutamic acid; F, phenylalanine; G, glycine; H, histidine; I, isoleucine; K, lysine; L, leucine; M, methionine; N, asparagine; P, proline; Q, glutamine; R, arginine; S, serine; T, threonine; V, valine; W, tryptophan; and Y, tyrosine.

terminal acetylation (*7*). Although all are glycoproteins, their widely scattered putative carbohydrate attachment sites occur at nonequivalent locations. All four sequences were analyzed by computer programs that predict secondary structure (*10*). Within the limitations of this approach, the distribution of α-helical segments and β-sheet sections was found to be reasonably similar. Similarly, the hydropathy profiles (*11*) of all four sequences were found to be equally undistinguished and to be characteristic of soluble globular proteins.

The four sequences were optimally aligned by pairs, and the statistical significance of the alignment scores was determined (Table 1). Angiotensinogen most resembles α_1-antitrypsin and is least similar to antithrombin III. The similarities are less than were observed previously for the antitrypsin-ovalbumin-antithrombin triad, but they are still highly significant in two cases (Table 1). Comparisons of alignments among the four sequences (Fig. 1) reveal that "invariant" residues occur at 31 of the aligned positions in the 400-residue bloc that spans them. At 27 other positions, angiotensinogen is identical to two of the other three, and at an additional 63 places it matches one of the three. Thus, at more than a third of the positions angiotensinogen has a residue in common with at least one of the other three proteins. The case for common ancestry is, in my view, indisputable.

On the basis of distantly related proteins whose primary and tertiary structures are both known, it can be expected that the three-dimensional resemblance of angiotensinogen and the protease-inhibitor group will be strong. In the case of serine proteases, for example, comparisons of α-carbon backbone coordinates are about twice as effective as sequence matching in detecting homology (*12*).

The evolution of a biologically active peptide as a portion of a protein of large molecular weight is of considerable interest. Although not enough data are available to make a judgment about which of the four related proteins appeared first, a possible scenario could involve chance mutations in some protease inhibitor that rendered the protein susceptible to cleavage near its amino terminus. If the amino-terminal segment so removed was not essential for the inhibitor function, then the gradual generation of a biologically active constellation of amino acids, comprising the removable portion, does not seem unreasonable, as long as a suitable receptor for the peptide existed. Alternatively, the biologically active peptide could conceivably have evolved as the result of an internal duplication of a signal peptide. It is perhaps noteworthy that the angiotensinogen segment destined to be released as the biologically active angiotensins occurs at a point which is proximal to the initiation points of two of the other three proteins, but which overlaps the signal peptide of the third (Fig. 1). In any event, the question arises as to whether the present protein—before or after the release of angiotensin—is a protease inhibitor, and, if so, what proteases does it regulate?

RUSSELL F. DOOLITTLE
Department of Chemistry, University of California, San Diego, La Jolla 92093

References and Notes

1. R. F. Doolittle, *Science* **214**, 149 (1981). The database searched included about 800 sequences collected since 1978 (NEWAT) and 1081 sequences from the 1978 version of the *Atlas of Protein Sequence and Structure* (National Biomedical Research Foundation, Washington, D.C., 1978), vol. 5, Suppl. 1–3.
2. H. Ohkubo, R. Kageyama, M. Ujihara, T. Hirose, S. Inayama, S. Nakanishi, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 2196 (1983).
3. K. Kurachi *et al.*, *ibid.* **78**, 6826 (1981).
4. R. W. Carrell *et al.*, *Nature (London)* **298**, 329 (1982).
5. T. E. Peterson, G. Dudek-Wojciechowska, L. Sotterup-Jensen, S. Magnusson, in *The Physiological Inhibitors of Blood Coagulation and Fibrinolysis*, D. Collen, B. Wiman, M. Verstraete, Eds. (Elsevier, Amsterdam, 1979), pp. 43–54.
6. L. T. Hunt and M. O. Dayhoff, *Biochem. Biophys. Res. Commun.* **95**, 864 (1980).
7. L. McReynolds *et al. Nature (London)* **273**, 723 (1978).
8. T. Chandra, R. Stackhouse, V. J. Kidd, S. L. C. Woo, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 1845 (1983).
9. S. H. Snyder and R. B. Innis, *Annu. Rev. Biochem.* **48**, 755 (1979).
10. P. Y. Chou and G. D. Fasman, *ibid.* **47**, 251 (1978); J. Garnier, D. J. Osgothorpe, B. Robson, *J. Mol. Biol.* **120**, 97 (1978).
11. J. Kyte and R. F. Doolittle, *J. Mol. Biol.* **157**, 105 (1982).
12. M. N. G. James, L. T. J. Delbaere, G. D. Brayer, *Can. J. Biochem.* **56**, 396 (1978).
13. Optimal alignments of pairs of sequences were obtained with a program that uses a modification of the Needleman-Wunsch algorithm [S. B. Needleman and C. D. Wunsch, *J. Mol. Biol.* **48**, 443 (1970)]. The gap penalty was fixed at 2.5 times the value awarded for matched identities (*1*). The significance of alignment scores was determined by comparing multiple randomized sequences of the same lengths and compositions. Situations in which the authentic comparisons gave scores > 3.0 standard deviations above the mean jumble comparisons were regarded as implying common ancestry.
14. I thank Karen Anderson for assistance and Harvey Itano for reading the manuscript and offering helpful suggestions. Supported in part by NIH grant RR 00757.

17 May 1983; accepted 17 September 1983

Male Esterase 6 Catalyzes the Synthesis of a Sex Pheromone in *Drosophila melanogaster* Females

Abstract. *Esterase 6, a component of the seminal fluid of* Drosophila melanogaster *males, hydrolyzes* cis*-vaccenyl acetate, a lipid made only by males, to* cis*-vaccenyl alcohol. This reaction occurs in the female reproductive tract and is virtually complete within 6 hours after copulation. Both the alcohol and the acetate decrease the number of matings among pairs of virgin flies in which the female is treated topically with these substances. Although females tested 10 minutes after copulation elicit less courtship than virgin females, females tested 6 hours after copulation stimulate even less courtship if they received active esterase 6 in the seminal fluid of their respective mates. Either the alcohol or a derivative appears to be an antiaphrodisiac that decreases courtship elicited by inseminated females and thus reduces the probability of further mating. Thus the activity of the pheromone depends on a final reaction which occurs in the female, using both substrate and enzyme provided by the male.*

Esterase 6 (EST 6) is a carboxylesterase (E.C. 3.1.1.1) of *Drosophila melanogaster*, and its locus (*Est-6*) is polymorphic for two or more alleles in most natural populations (*1*). The enzyme is concentrated in the anterior ejaculatory duct (*2*), a secretory and propulsive organ of the male reproductive system, and is transferred to females within 2 to 3 minutes of the initiation of copulation (*3*). The presence of EST 6 in the male ejaculate is associated with effects on sperm utilization, female fertility (progeny production at 18°C), and the timing of female remating (*4*).

An acetate ester, *cis*-vaccenyl acetate (cVA), is found in the male ejaculatory bulb of the mature adult and is transferred to females during mating (*5*). The presence of cVA in the female reproductive tract may inhibit or prevent further courtship of inseminated females by acting as an antiaphrodisiac pheromone (*6*). The close proximity of EST 6 and cVA in the male reproductive system, the transfer of the two compounds to the female during copulation, and the chemical structure of cVA suggested that it might be an in vivo substrate for EST 6 in the female reproductive tract.

A mixture of the electrophoretically fast and slow allozymes of EST 6 purified from an Ore R strain of *Drosophila melanogaster* (*7*) was incubated with