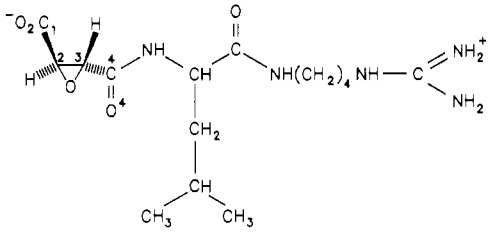
# Crystal Structure of a Papain–E-64 Complex†

K. I. Varughese,‡ F. R. Ahmed, P. R. Carey, S. Hasnain, C. P. Huber,* and A. C. Storer§
*Division of Biological Sciences, National Research Council of Canada, Ottawa, Ontario K1A 0R6, Canada*
*Received May 24, 1988; Revised Manuscript Received September 12, 1988*

ABSTRACT: E-64 [1-[*N*-[(L-3-*trans*-carboxyoxirane-2-carbonyl)-L-leucyl]amino]-4-guanidinobutane] is an irreversible inhibitor of many cysteine proteases. A papain–E-64 complex was crystallized at pH 6.3 by using the hanging drop method. Three different crystal forms grew in 3–7 days; the form chosen for structure analysis has space group $P2_12_12_1$, with $a$ = 42.91(4) Å, $b$ = 102.02(6) Å, $c$ = 49.73(2) Å, and $Z$ = 4. Diffraction data were measured to 2.4-Å resolution, giving 9367 unique reflections. The papain structure was solved by use of the molecular replacement method, and then the inhibitor was located from a difference electron density map and fitted with the aid of a PS330 computer graphics system. The structure of the complex was refined to $R$ = 23.3%. Our analysis shows that a covalent link is formed between the sulfur of the active-site cysteine 25 and the C-2 atom of the inhibitor. Contrary to earlier predictions, the E-64 inhibitor clearly interacts with the S subsites on the enzyme rather than the S′ subsites, and papain's histidine 159 imidazole group plays a binding rather than a catalytic role in the inactivation process.

The epoxide 1-[*N*-[(L-3-*trans*-carboxyoxirane-2-carbonyl)-L-leucyl]amino]-4-guanidinobutane (I), also known



I

as E-64, was first isolated from a culture of *Aspergillus japonicus*, and its structure was determined by Hanada et al. (1978a,b). Since then it has been shown that E-64 is a potent irreversible inhibitor of several cysteine proteases, e.g., papain, ficin, cathepsin L (Towatari et al., 1978), cathepsin B (Inaba et al., 1979), and calpain (Sugita et al., 1980). The potency of the inhibitor, coupled with its low toxicity, indicates a potential use for the epoxide and its derivatives as a drug for the suppression of elevated levels of cysteine protease activity associated with certain disease states, e.g., muscular dystrophy (Kar & Pearson, 1976; Pennington & Robinson, 1968). Thus, a knowledge of the mechanism of inhibition and of the enzyme inhibitor interactions is very valuable for the design of E-64 analogues with increased potency and selectivity.

The structure of E-64 has been shown to contain a *trans*-epoxysuccinic acid attached to a modified dipeptide [(leucylamino)-4-guanidinobutane] (Hanada et al., 1978b). Prior to the present study the mechanism of inactivation of cysteine proteases by E-64 was not known. The site of alkylation was inferred to be the active-site residue cysteine 25 (Hanada et al., 1978b). Barrett et al. (1982) have proposed that the dipeptide unit of E-64 binds in the leaving group side (S′) of the catalytic site. In a recent review Rich (1986) has suggested that the leucyl side chain binds at the $S'_2$ subsite and the amino-4-guanidinobutane binds at the $S'_3$ subsite, thus permitting attack of the cysteine 25 thiolate anion at the C-2 carbon of E-64 assisted by the protonation of the epoxide oxygen by the imidazolium group of histidine 159. The structure described below clearly shows that while the covalent bond is between the cysteine 25 sulfur and the C-2 carbon of E-64, the binding of E-64 with papain, and most probably cysteine proteases in general, is not as previously proposed.

## Experimental Procedures

Fully active papain was prepared as described by Blumberg et al. (1970) from the twice-crystallized enzyme purchased from Sigma Chemical Co. The purified enzyme was concentrated to 5.2 mg/mL, and 1.6 equiv of E-64 was added. The inhibited enzyme was dialyzed against 20% methanol, and 10 mM sodium phosphate buffer, pH 6.3, was then added. For crystallization the reservoir contained 500 μL of $H_2O$/methanol (1:2) and 60 μL of saturated NaCl solution. Crystals of the complex were obtained in 3–4 days by the hanging drop method using Linbro plates. Three different forms of crystals were obtained; two were found to be stable, and the other was unstable and could not be mounted. The first form was orthorhombic, $P2_12_12_1$, $a$ = 42.91(4) Å, $b$ = 102.02(6) Å, $c$ = 49.73(2) Å, and $Z$ = 4. The second was tetragonal with a very large cell, $a$ = 145 Å and $c$ = 256 Å. The orthorhombic form was chosen for structure analysis, as it diffracted well and contained only one molecule per asymmetric unit. This form is different from any of the six crystal forms observed by Drenth et al. (1971). The crystal used for intensity data measurement was 0.3 × 0.3 × 0.2 mm, and the intensity data were collected by using an Enraf-Nonius CAD4 diffractometer with Ni-filtered Cu Kα radiation. A total of 9367 unique reflections to a resolution limit of 2.4 Å were measured, of which 6467 had $F_o \geq 2\sigma(F)$. The three intensity control reflections decreased by 22% toward the end of the data collection. A decay correction was applied as a function of time. The ω-scan technique was employed, $\Delta\omega = (0.6 + 1.4 \tan\theta)°$ for the peak and 25% on each side for the background. The reflections were scanned at the rate of 1°/min for $\theta \leq 15°$ and 0.67°/min for $\theta > 15°$.

*Structure Solution.* The structure was solved by use of molecular replacement techniques. A program package (B.

† NRCC Publication No. 29779.
* Author to whom correspondence should be addressed.
‡ Present address: Department of Biology, University of California, San Diego, La Jolla, CA 92093.
§ Present address: Biotechnology Research Institute, National Research Council of Canada, 6100 Royalmount Avenue, Montréal, Québec H4P 2R2, Canada.

0006-2960/89/0428-1330$01.50/0 Published 1989 by the American Chemical Society

M. Craven, Technical Report TR-75-2, 1975) consisting of Crowther's fast rotation (Crowther, 1969) and Langs' translation function (Langs, 1975) was used for the structure solution. The search model consisted of 1655 non-hydrogen atoms of papain refined with 1.65-Å resolution data (Kamphuis et al., 1984). The search model was put into a triclinic cell of size 55 × 70 × 55 Å with $\alpha = \beta = \gamma = 90°$, and then $(E^2 - 1)$ values were computed and used as Patterson coefficients. There were 1548 reflections with $(E^2 - 1) > 0.0$ in the resolution range 5–25 Å. For the observed Patterson, 1937 reflections in the same resolution range were used. $E^2$ values modified by $\exp(-10.0 \sin^2 \theta/\lambda^2)$ were used as coefficients to minimize the effects of series termination. The highest peak found in the rotation search, at Eulerian angles $\alpha = 6°$, $\beta = 7.2°$, and $\gamma = 215°$, proved to be correct. The height of the next highest peak was about two-thirds of the first. Translation search was carried out by using the function defined by Langs (1975), which is basically a Patterson function of the deconvoluted molecular structure. The translation components were located from the Harker sections $x = {}^1/_2$, $y = {}^1/_2$, and $z = {}^1/_2$. The peaks giving the correct solution were at least twice as high as the second highest peak in each section.

*Refinement and Location of the Inhibitor.* Inhibitor fitting was performed by using a $(F_o - F_c)$ map with the FRODO program (Jones, 1978; Pflugrath et al., 1984) on an Evans and Sutherland PS330 interactive computer graphics system. Initially it was not known how the E-64 was attached to papain. However, the electron density distribution clearly indicated that a covalent bond was formed between the C-2 carbon of the inhibitor and the sulfur of the papain residue cysteine 25.

Six cycles of restrained refinement were carried out after the initial model for the inhibitor was included. The inhibitor model was then refitted with the help of a $(F_o - F_c)$ map with the inhibitor atoms omitted from the structure factor calculation. The location of the 4-guanidinobutane portion of E-64 was indicated in the electron density distribution only after 14 cycles of refinement. For these first 14 cycles an overall temperature factor ($B = 11.0$ Å$^2$) was used; subsequently, individual isotropic thermal parameters were refined. Following four more cycles it was noted that arginine 93 was making short contacts with symmetry-related molecules, and the side-chain position was adjusted by using a $(3F_o - 2F_c)$ map. Water molecules were also located at this point. After several more sequences of least-squares refinement cycles and refitting, the refinement terminated with an $R$ value of 0.233 for the 6244 reflections with $F_o > 1.4\ \sigma(F_o)$ and with $\sin \theta/\lambda > 0.09$. We used the programs written by Hendrickson and Konnert (1980) for restrained least-squares refinement.

## Results and Discussion

The active site of papain has been proposed by Schechter and Berger (1967) to consist of seven subsites: $S_4$ to $S_1$ bind the four peptidyl substrate residues immediately to the N-terminal side of the scissile bond, and $S'_1$ to $S'_3$ bind the three residues immediately to the C-terminal side. It was previously proposed (Barrett et al., 1982; Rich, 1986) that, unlike chloromethyl ketone inhibitors which bind in the S subsites (Drenth et al., 1976), E-64 binds in the S′ subsites. However, a schematic comparison (Figures 1 and 2) of the interactions of benzyloxycarbonylphenylalanylalanine chloromethyl ketone (BPACK) (Drenth et al., 1976) and of E-64 with papain clearly shows that both inhibitors interact with the same subsites on the enzyme and that the interactions utilized are very similar in each case. In the papain E-64 complex the O-4 carbonyl of the E-64 forms a H-bond to the NH of glycine

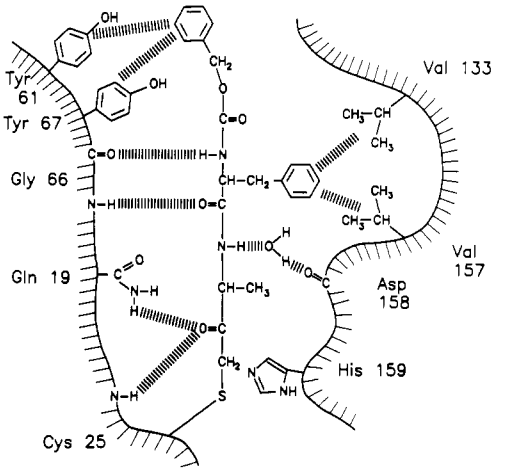



FIGURE 1: Schematic diagram showing the interactions of benzyloxycarbonylphenylalanylalanine chloromethyl ketone inhibitor with papain.

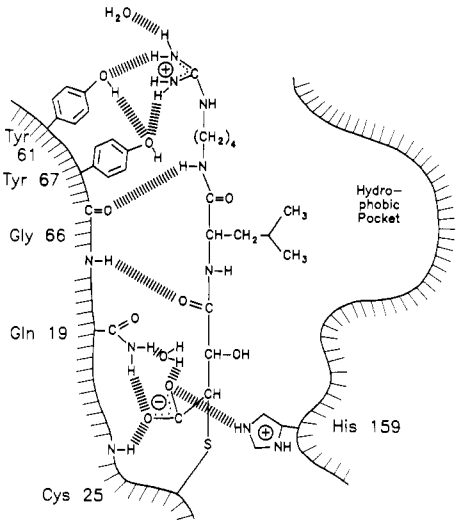



FIGURE 2: Schematic diagram showing the interactions of E-64 with papain.

Table I: Hydrogen Bonds Involving E-64 and Papain

| E-64 | papain | distance (Å) | E-64 | papain | distance (Å) |
|---|---|---|---|---|---|
| O1 | ND1(159) | 2.90 | N2 | O(66) | 3.04 |
| O1 | O(217) | 2.67 | N4 | OH(61) | 2.96 |
| O2 | N(25) | 2.95 | N4 | O(216) | 2.45 |
| O2 | NE2(19) | 2.87 | N5 | OH(67) | 2.93 |
| O4 | N(66) | 2.88 | | | |

66 while N-2 of E-64 binds to the carbonyl oxygen of glycine 66. Table I lists the hydrogen-bond distances involving E-64. Thus, whereas in the chloromethyl ketone structure, determined by Drenth et al. (1976), the phenylalanine of the inhibitor bound in the $S_2$ subsite forms a one-residue stretch of antiparallel β-sheet with glycine 66, the equivalent interactions between E-64 and glycine 66 constitute a one-residue stretch of parallel β-sheet. The E-64 leucyl side chain is located at the entry to the hydrophobic pocket of papain's $S_2$ subsite but does not extend nearly as far into the pocket as the phenylalanine side chain of the chloromethyl ketone inhibitor (Drenth et al., 1976). In the current structure the 4-guanidinobutane moiety of E-64 is less well defined than the rest of the inhibitor but appears to occupy a position similar to the benzyloxy group in the BPACK complex with papain. In the present case one of the terminal amino groups of the guanidino portion is hydrogen-bonded to the terminal hydroxyl oxygen of tyrosine 61 and to a water molecule, while the other amino group forms

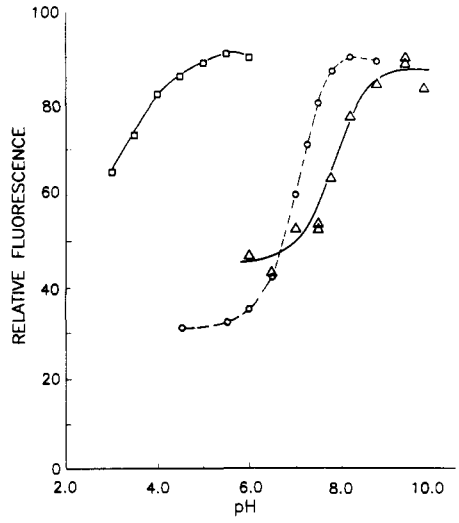



FIGURE 3: Fluorescence titration of active papain (Δ), mercury papain (□), and papain–E-64 complex (○) at 20 °C. The excitation wavelength was 290 nm, and the emission wavelength was 330 nm. Buffers used were as follows: pH 3.0–5.5, 0.1 M acetate and 0.2 M NaCl; pH 6.0–7.5, 0.1 M phosphate and 0.1 M NaCl; pH 8.0–9.0, 0.1 M borate and 0.1 M NaCl. In all cases enzyme concentration was 3 μM.

a hydrogen bond to the hydroxyl group of tyrosine 67. The carboxylic acid group of E-64 participates in a complex network of H-bonds involving several groups on the enzyme. One of the carboxylic oxygens occupies the oxyanion hole first described by Drenth et al. (1976) and forms H-bonds to the side chain of glutamine 19 and to the backbone NH of cysteine 25. The second carboxylic oxygen also participates in two H-bonds, one to a water molecule that is, in turn, H-bonded to the side chain of glutamine 19 [$H_2O$···NE2(19) = 2.95 Å] and the second to the imidazole ring of histidine 159. This second H-bond requires that either the carboxylic acid of E-64 or the histidine 159 of the enzyme be protonated. The second option is more likely at the pH of crystallization (i.e., pH 6.3), and evidence supporting this is shown in Figure 3. The fluorescence intensity of tryptophan 177 has been shown to be influenced by the ionization state of histidine 159 (Johnson et al., 1981); i.e., the protonation of histidine 159 quenches the fluorescence of tryptophan 177. This property of papain was used to determine a $pK_a$ of 7.0 for the side chain of histidine 159 in the papain–E-64 complex (Figure 3). On the basis of the current X-ray structural information for the papain–E-64 complex, the most likely mechanism of formation of the covalent bond between the inhibitor and the enzyme involves the attack of the sulfur of cysteine 25 on the C-2 carbon of E-64. Protonation of the epoxide oxygen must involve water rather than the imidazole of histidine 159 as previously suggested (Rich, 1986). This mechanism is supported by the fact that the protonated imidazole is occupied in a H-bond with the carboxylate group of E-64, and also the resultant hydroxyl group of the adduct points away from histidine 159 into the solvent. Formation of the carbon–sulfur bond, which is not easily broken, accounts for the irreversible nature of the complex.

Comparison of the enzyme structure in the E-64 complex with the native papain structure (Kamphuis et al., 1984) indicates a slight widening of the active-site cleft in this complex. The backbone in the cysteine 63 to tyrosine 67 region is shifted by about 1.2 Å away from the inhibitor, but this is partially compensated for by a shift of about 0.5 Å toward the inhibitor by the backbone in the aspartic acid 158–alanine 160 region on the other side of the cleft. The peptide loop of which serine 21 is the end is also shifted by about 1 Å toward the asparagine 64 loop. In this structure, in contrast to the BPACK complex, the side chain of glutamine 19 is rotated by approximately 60° around the CA–CB bond and approximately 120° around the CG–CD bond, relative to the native conformation, in order to facilitate the various hydrogen bonds. With this change, it becomes unnecessary to rotate the side chain of serine 176 in order to maintain the OH(176)···OE1(19) hydrogen bond (O···O distance 2.81 Å). The other obvious difference in the enzyme portion of E-64 complex is the orientation of the arginine 93 side chain, already mentioned. This last difference is clearly due to the different packing in the two unit cells.

## ACKNOWLEDGMENTS

We are grateful to Professor Jan Drenth for providing the refined coordinates of papain prior to publication.

**Registry No.** L-Cys, 52-90-4.

## REFERENCES

Barrett, A. J., Kembhavi, A. A., Brown, M. A., Kirschke, H., Knight, C. G., Tamai, M., & Hanada, K. (1982) *Biochem. J. 201*, 189–198.

Blumberg, S., Schechter, I., & Berger, A. (1970) *Eur. J. Biochem. 15*, 97–102.

Crowther, R. A. (1969) *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem. B25*, 2571–2580.

Drenth, J., Jansonius, J. N., Koekoek, R., & Wolthers, B. G. (1971) *Adv. Protein Chem. 25*, 79–115.

Drenth, J., Kalk, K. H., & Swen, H. M. (1976) *Biochemistry 15*, 3731–3738.

Hanada, K., Tamai, M., Yamagishi, M., Ohmura, S., Sawada, J., & Tanaka, I. (1978a) *Agric. Biol. Chem. 42*, 523–528.

Hanada, K., Tamai, M., Ohmura, S., Sawada, J., Seki, T., & Tanaka, I. (1978b) *Agric. Biol. Chem. 42*, 529–536.

Hendrickson, W. A., & Konnert, J. H. (1980) in *Computing in Crystallography* (Diamond, R., Ramaseshan, S., & Venkatesan, K., Eds.) pp 13.01–13.23, Indian Institute of Science, Bangalore.

Inaba, T., Hirayama, Y., & Fujinaga, N. (1979) *Agric. Biol. Chem. 43*, 655–656.

Johnson, F. A., Lewis, S. D., & Shafer, J. A. (1981) *Biochemistry 20*, 52–58.

Jones, T. A. (1978) *J. Appl. Crystallogr. 11*, 268–272.

Kamphuis, I. G., Kalk, K. H., Swarte, M. B. A., & Drenth, J. (1984) *J. Mol. Biol. 179*, 233–256.

Kar, N. C., & Pearson, C. M. (1976) *Clin. Chim. Acta 73*, 293–297.

Langs, D. A. (1975) *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr. A31*, 543–550.

Pennington, R. J., & Robinson, J. E. (1968) *Enzymol. Biol. Clin. 9*, 175–182.

Pflugrath, J. W., Saper, M. A., & Quiocho, F. A. (1984) in *Methods and Applications in Crystallographic Computing* (Hall, S. R., & Ashida, T., Eds.) pp 404–407, Clarendon, Oxford, U.K.

Rich, D. H. (1986) in *Proteinase Inhibitors* (Barrett & Salvesen, Eds.) pp 153–178, Elsevier, Amsterdam.

Schechter, I., & Berger, A. (1967) *Biochem. Biophys. Res. Commun. 27*, 157–162.

Sugita, H., Ishiura, S., Suzuki, K., & Imaghari, K. (1980) *J. Biochem. (Tokyo) 87*, 339–341.

Towatari, T., Tanaka, K., Yoshikawa, D., & Kataunuma, N. (1978) *J. Biochem. (Tokyo) 84*, 659–671.