THE EVOLUTION OF PAPAIN

Boris Weinstein

Department of Chemistry, University of Washington
Seattle, Washington 98105

Received September 14, 1970

Summary. Papain possesses several regions of internal homology and similar shape. The enzyme probably formed through a series of gene doublings.

Archetype correlations between glucagon-secretin (Weinstein, 1968) and rubredoxin-ferredoxin (Weinstein, 1969) have been based on the process of gene duplication. Attention is now directed to the plant proteinase, papain. Recently, an x-ray study of this compound revealed the polypeptide chain to be folded into two distinct parts, separated by a groove (Drenth, et al, 1968). On either side of the depression lie a cysteine and a histidine. These two residues apparently constitute the active site of the enzyme. Further, the investigation showed that the earlier and tentative amino acid arrangement needed considerable modification (Light, et al, 1964). As a result, two groups have presented revised data on the primary sequence (Husain, et al, 1969; Mitchel, et al, 1970).

The symmetrical shape of papain, divisible at about residue 110, suggested a search for internal homologies between these regions. Segment 27-50 is duplicated both at 160-183 and 185-199, while 8-18 is again seen at 122-132 and 202-212. A third, tentative correlation lies between 48-54 and 116-122, where short pieces of α-helix are present. Other minor duplications exist, but they appear inconclusive on statistical grounds (Haber, et al, 1970). A few additional alignments become available through the introduction of gaps; however, the lack of variant papains for comparison does not encourage this approach. The principle relationships are summar-

```
27                                  36                          43                      50
Ala Phe Ser Ala Val Val Thr Ile Glu Gly Ile Ile Lys Ile Arg Thr Gly Asn Leu Asn Gln Tyr Ser Glu

160                                 169                         176                     183
Ala Val Ala Ala Val Gly Tyr Asn Pro Gly Tyr Ile Leu Ile Lys Asn Ser Trp Gly Thr Gly Trp Gly Glu

                                    185                         192                     199
                                    Gly Tyr Ile Arg Ile Lys Arg Gly Thr Gly Asn Ser Tyr Gly Val


 8                                      18                      48                  54
Arg Gln Lys Gly Ala Val Thr Pro Val Lys Asn                     Tyr Ser Glu Gln Glu Leu Leu

122                                     132                     116                 122
Leu Tyr Ser Ile Ala Asn Gln Pro Val Ser Val                     Tyr Asn Gln Gly Ala Leu Leu

202                                      212
Leu Tyr Thr Ser Ser Phe Tyr Pro Val Lys Asn
```

Figure 1. Long internal homologies in papain. Identical positional residues are underscored.

ized in Figure 1.

It is interesting to note that the first half of the protein has two miniature wings, whose walls are constructed by sections 69-81 and 52-65, plus 95-100 (Dickerson, et al, 1969). The cystine bridges between 22-63 and 56-95 are included within this area. Juxtaposed in space to cysteine 56 is histidine 81. This pair may mark the original sulfhydryl active site in a more primitive papain.

These relationships suggest the ancestral precursor began as a protein with about 55-65 residues, followed by a doubling to 110-130 units. A reduplication then extended the chain to about 185 amino acids. Later, two shorter extensions afforded the present day papain. It is realized such arguments are speculative in nature, but they can serve as a stimulus to the future synthesis of modified or new papains. Certainly, the availability of other sulfhydryl plant proteinases would act both as a guide and an extension to additional work in this area.

ACKNOWLEDGEMENT

The author wishes to thank the National Institute of Health for partial support (AM 12616).

REFERENCES

Dickerson, R. E., and Geis, I., Stereo Supplement to The Structure and Action of Proteins, Harper and Row, New York, 1969, p. 18.

Drenth, J., Jansonius, J. N., Koekoek, R., Swen, H. M., and Wolthens, B. G., Nature, 218, 929 (1968).

Haber, J. E., and Koshland, Jr., D. E., J. Mol. Biol., 50, 617 (1970).

Hasain, S. S., and Lowe, G., Biochem. J., 114, 279 (1969).

Light, A., Frater, R., Kimmel, J. R., and Smith, E. L., Proc. Nat. Acad. Sci. U. S. A., 52, 1276 (1964).

Mitchel, R. E. J., Chaiken, I. M., and Smith, E. L., J. Biol. Chem., 245, 3485 (1970).

Weinstein, B., Experientia, 24, 406 (1968).

Weinstein, B., Biochem. Biophys. Res. Commun., 35, 109 (1969).