Subtilisin Carlsberg

V. THE COMPLETE SEQUENCE; COMPARISON WITH SUBTILISIN BPN'; EVOLUTIONARY RELATIONSHIPS*

(Received for publication, November 14, 1967)

EMIL L. SMITH, ROBERT J. DELANGE, W. HOWARD EVANS, MICHAEL LANDON, AND FRANCIS S. MARKLAND

From the Department of Biological Chemistry, School of Medicine, and the Molecular Biology Institute, University of California, Los Angeles, California 90024

SUMMARY

Evidence is presented for the complete amino acid sequence of subtilisin Carlsberg. The protein consists of a single peptide chain of 274 residues. Comparison with subtilisin BPN' shows 84 amino acid differences and 1 additional residue in BPN'. The 84 differences can be accounted for on the basis of single or double nucleotide replacements. Within the subtilisins, there are a number of distinct repetitions of sequence; this suggests that the proteins may have evolved from shorter peptide chains by some process of extension of the sequence. The subtilisins differ in structure from other proteinases sensitive to diisopropyl fluorophosphate, as well as from proteinases of other types.

In the preceding papers of this series on the amino acid sequence of subtilisin Carlsberg, the isolation and sequences of the peptides from tryptic (2, 3) and chymotryptic digests have been reported (4, 5). In this paper we shall demonstrate that sufficient information is available to establish the complete sequence unequivocally. Further discussion will be concerned with a comparison of this sequence with that determined for subtilisin BPN' (6), and the possible mutational events that led to the differences in the primary structures of these enzymes. Segments of the sequence of possible significance in the function or conformation of these enzymes will also be considered.

PRIMARY STRUCTURE OF SUBTILISIN CARLSBERG

For convenience we shall review briefly the results described in the preceding papers. Subtilisin Carlsberg contains only 9 lysine and 4 arginine residues, indicating that only 14 tryptic peptides are expected. All 14 of these peptides were isolated in pure form, and their compositions account for 274 residues, in excellent agreement with the 276 residues that were estimated from the analysis of the intact protein (2). The complete sequences of nine of these peptides were elucidated: Peptides T-1, T-2, T-3, T-4, T-7, T-8, T-11, T-12, and T-13 (3). Partial sequences of the remaining peptides were also determined, and, in all cases, a sufficient number of residues was established at the amino and carboxyl ends to place the peptides unequivocally. Furthermore, other portions of the sequences of Peptides T-5, T-6, T-10, and T-14 were elucidated to eliminate ambiguities because of insufficient overlaps with the chymotryptic peptides.

The chymotryptic peptides accounted for 272 of the 274 residues (4, 5). In addition, chymotryptic peptides were isolated that represented incompletely hydrolyzed segments, i.e. C-4 + 5, C-8 + 9, C-9 + 10, C-17 + 18, C-25 + 26, and C-32 + 33; these were particularly valuable in placing chymotryptic peptides within the larger tryptic peptides.

Since the complete sequence has been established, both the tryptic and chymotryptic peptides are numbered in the order in which they occur in the sequence, beginning with Peptides T-1 and C-1, respectively, at the amino terminus, and ending with Peptides T-14 and C-33 at the carboxyl terminus of the protein.

Fig. 1, A and B, presents the complete sequence of subtilisin Carlsberg and the major peptides used in deducing this sequence. We shall discuss only those critical features that are not readily apparent from inspection of the figure.

The position of Peptide T-1 at the amino terminus is established from the earlier work of Ottesen and Schellman (7), who showed that Ala-Glx is at the NH₂ terminus of the protein, and from our finding that only Peptide T-1 possesses this sequence (Fig. 1A). Peptide T-14 is the only major tryptic peptide lacking a basic residue (8) and thus must be derived from the COOH-terminal segment of the protein (Fig. 1B). In addition, this peptide possesses essentially the same sequence as Peptide T-14 of subtilisin BPN', which is known to be COOH-terminal in the homologous protein (8).

The evidence for the positioning of Peptides T-2 and T-3...
following T-1 is evident. The order of Peptides T-4 and T-5 is
given by the composition of Peptide C-4 (Fig. 1A). The se-
quence of Peptide T-9 (residues 28 through 94) was determined
mainly from the chymotryptic peptides derived from the entire
protein (5). Proof that Peptides C-5 through C-11 were derived
from the segment represented by Peptide T-5 is available from
the isolation of peptides of identical compositions obtained from
a chymotryptic digest of Peptide T-5 (3). Note that residue
56 is omitted in order to preserve the homologous numbering of
the residues with those of subtilisin BPN' (see below).

Within the sequence of Peptide T-5, no overlaps were available
from the chymotryptic digest for residues 50 and 51, for residues
57 and 58, and for residues 67 and 68. For residues 67 and 68,
the composition of Peptide T-5-C-6 (3), containing the 2 histidine
residues of Peptide T-5 and the sum of the compositions of
Peptides C-7 and C-8, is sufficient to show that Peptides C-7 and
C-8 are joined in that order. In order to establish conclusively
the position of Peptide C-6 and to be certain that no residues were
overlooked, Peptide T-5 was hydrolyzed with papain and
certain fractions were carefully examined (3). Key peptides
were isolated that closed the remaining gaps. The composition
of Peptide T-5-Pa-1, containing the sole phenylalanine residue
of Peptide T-5, provided evidence to position residues 50 and 51.
Similarly, Peptide T-5-Pa-2, a tyrosine-containing peptide of
unique composition, served to place residues 58 and 59. For
both these peptides derived by papain digestion, the points of
hydrolysis (Fig. 1A) are in accord with the specificity of this
enzyme.

The relative positions of Peptides T-5 and T-6 were established
by the unique sequence of Peptide C-11. The sequence of
Peptide T-6 was determined except for two short regions (3) that
were provided by Peptides C-12 and C-13, the latter containing
the sole tryptophan residue of the protein.

The relationships of Peptides T-6, T-7, T-8, and T-9 warrant
brief comment. For Peptides T-8 and T-9, the overlap consists
of only a single residue. Fortunately, Peptide T-9 is the only
tryptic peptide bearing NH₂-terminal tyrosine (3). For the
positioning of Peptides T-0 and T-7, Peptide C-10 supplies only
a single overlapping lysine residue, a situation similar to that
for Peptides T-3 and T-4. Since the overlaps for all other
lysine-containing peptides can be established by unambiguous
overlap of 2 or more residues, it is evident that Peptide T-3
cannot join to Peptide T-7, since this would leave no room in the
protein for the sequence represented by the three peptides which
are joined in the order (T-4)-(T-5)-(T-0).

The sequence of Peptide T-10 was established mainly by
chymotryptic peptides (5); however, the positioning of these
chymotryptic peptides was determined from fragments ob-
tained by cyanogen bromide cleavage of Peptide T-10 (3).
Peptide T-10-BrCN-1 placed Peptides C-20 and C-21 at the
NH₂ terminus. Peptide T-10-BrCN-3 possessed a composition
accounting for part of Peptide C-21 and all of Peptides C-22,
C-23, and C-24. This also placed Peptide C-24 at the COOH
end of Peptide T-10-BrCN-3, since methionine or homoserine is
at the COOH terminus of both. Peptide T-10-Pa-1 placed
Peptide C-22 COOH-terminal to Peptide C-21 (3). The other
cyanogen bromide peptides served to confirm the positions of
the remaining chymotryptic peptides and to establish the COOH
terminus of Peptide T-10 as Ile-Leu-Ser-Lys (3). Only the
dipeptide sequence Ile-Leu was not recovered from the chymo-
tryptic digest of the protein (5).

The sequences and relationships of the remainder of the
sequence of subtilisin Carlsberg are straightforward. Peptides
T-10 and T-11 are uniquely placed in that order, since the only
tryptic peptide with NH₂-terminal histidine is Peptide T-11.
The joining of Peptides T-13 and T-14 is established by Peptide
C-32. The dipeptide T-12, Asn–Arg, fits securely in the
sequence, since it is the only tryptic peptide with NH₂-terminal
asparagine, and Peptide C-29 overlaps with Peptide T-13, the
sole tryptic peptide with NH₂-terminal leucine.

The positioning and overlapping of the tryptic and chymo-
tryptic peptides were facilitated by the presence of small numbers
of arginine, phenylalanine, methionine, histidine, and trypto-
phan residues. Conversely, the large numbers of serine, glycine,
alanine, valine, and aspartic acid (or asparagine) residues
rendered the sequence determinations more difficult. These
five types of residues account for 61% of the protein and proved
to be, in some cases, heavily clustered in certain portions of the
sequence.

COMPARISON OF SUBTILISINS CARLSBERG AND BPN'

In Fig. 2, the complete sequences of the two presently estab-
lished subtilisins are shown for comparison. 1 The continuous
sequence is that of subtilisin BPN' (6) and the residues of
subtilisin Carlsberg that differ are shown above this sequence.
Since a residue is lacking in subtilisin Carlsberg, this is shown as
dash at residue 56. This position has been chosen in order to
minimize the apparent number of differences in the sequence as
well as to minimize the number of apparent mutational events
in the nucleotide codons represented in this region. Neverthe-
less, the deletion (or addition) of a residue must have occurred
between residues 50 and 59, in view of the excellent homology
of the sequences of the two subtilisins on either side of this por-
tion of the two proteins.

If the deletion (or addition) is excluded, the two subtilisins
differ in 84 residues. Table I gives the amino acid composition
of subtilisin BPN' and indicates the replacements in the Carls-
berg enzyme. From the presently assigned triplet nucleotide
codonts for the amino acids (9), 61 of the changes can be ascribed
to single base changes and 23 to double base changes. This is
on the basis of the minimal number of base changes required to
produce the amino acid substitution, ignoring codons that would
require the assumption of a greater number of mutational events.
The 274 residues present in both subtilisins differ in 30.6% of
the sequence. If we assume an equal rate of point mutation
(single base substitution) in the 84 altered codons, we should
expect 30.6% alterations, or 25.7 double base changes, equivalent
to 25.7 residues. (Mutations that do not result in an amino
acid substitution are not counted.) This is in good accord with
the 23 replacements that were actually found and can be as-
signed as double base changes. Although this does not prove
that the two homologous genes for these subtilisins evolved
from their common ancestral gene solely by the process of
accumulating point mutations, the data do suggest that no
special assumptions need be made, except for the loss (or gain)
of a triplet codon to explain the deletion (or addition) of residue
56.

1 The assignment of the amides of several of the residues in both
subtilisins differs from that presented earlier (1). The corrected
assignments are given for subtilisin Carlsberg in preceding papers
of this series (3, 5), and for subtilisin BPN' by Markland and
Smith (6).
FIG. 1A. Amino acid sequence of subtilisin Carlsberg for residues 1 through 152, inclusive. Proof of the sequence is derived from the overlapping tryptic (T) and chymotryptic (C) peptides. Continuous lines indicate that the exact positions of the residues were established; dashed lines indicate that only the compositions of the overlapping peptides were determined. The dash at residue 56 indicates that no residue is present, but the numbering has been kept the same as for subtilisin BPN'.

The number of changed residues of the different kinds is of some interest. The percentage changes given in Table I are, of course, meaningful only for those amino acid residues present in significant numbers. A striking feature of this tabulation is that only 12% of the glycine residues are altered. It may be recalled that, for the homologous cytochromes c from many species, more glycine residues are constant than any others (10, 11). The special importance of glycine for protein conformation has been discussed elsewhere (10).

Although the sequences of only two subtilisins are presently available, certain features of the amino acid substitutions are worth noting. Of the residues possessing large hydrophobic
FIG. 1B. Amino acid sequence of subtilisin Carlsberg for residues 153 through 275, inclusive. The asterisk on residue 221 indicates that this is the serine residue which is reactive with diisopropyl fluorophosphate.

Examples that illustrate this point are: serine and threonine (seven times); serine and asparagine (seven times); alanine and serine (five times); threonine and alanine (five times), etc. In addition, there are many replacements among hydrophilic residues of various types, as well as various radical replacements.

We assume that radical replacements involving side chains of very different size or chemical character are probably on the outside of the protein. The occurrence of tryptophan in the same position as a glycine (residue 106) serves as an example of such radical replacement.

A few additional replacements should be noted. Kinetic studies have indicated that at least 1 histidine residue undoubtedly is involved in the mechanism of action of the enzyme (13). Only the histidine at residue 17 in subtilisin BPN' can be excluded from consideration, since it is replaced by glutamine in...
FIG. 2. The continuous sequence is that of subtilisin BPN'. The residues which differ in subtilisin Carlsberg are given above the corresponding residue. The dash at residue 56 indicates that the corresponding residue is lacking in the Carlsberg enzyme.

**TABLE I**

**Amino acid substitutions in subtilisins**

<table>
<thead>
<tr>
<th>Residue</th>
<th>Subtilisin BPN</th>
<th>Changes</th>
<th>Subtilisin Carlsberg</th>
<th>Changed residue (No.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lysine</td>
<td>11</td>
<td>36%</td>
<td>Asn (2), Thr (1), Tyr (1)*</td>
<td>Asn (1), Ala (1)</td>
</tr>
<tr>
<td>Histidine</td>
<td>6</td>
<td>17%</td>
<td>Gln (1)</td>
<td></td>
</tr>
<tr>
<td>Arginine</td>
<td>2</td>
<td>0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aspartic acid</td>
<td>11</td>
<td>45%</td>
<td>Gln (1), Gly (1), Glu (1), Ser (2)</td>
<td></td>
</tr>
<tr>
<td>Asparagline</td>
<td>17</td>
<td>41%</td>
<td>Tyr (1), Asp (1), Ser (3), Gly (1), Ala (1)</td>
<td></td>
</tr>
<tr>
<td>Glutamic acid</td>
<td>4</td>
<td>25%</td>
<td>Ser (1)</td>
<td></td>
</tr>
<tr>
<td>Glutamine</td>
<td>11</td>
<td>64%</td>
<td>Leu (1), Thr (1), Ser (2), Asn (1), Tyr (1), Glu (1)</td>
<td></td>
</tr>
<tr>
<td>Threonine</td>
<td>13</td>
<td>46%</td>
<td>Lys (1), Ala (2), Ser (3)</td>
<td></td>
</tr>
<tr>
<td>Serine</td>
<td>37</td>
<td>49%</td>
<td>Thr (4), Pro (1), Ala (3), Gly (3), Val (1), Arg (2), Asn (4)</td>
<td></td>
</tr>
<tr>
<td>Proline</td>
<td>14</td>
<td>43%</td>
<td>Asp (2), Ala (3), deletion (1)</td>
<td></td>
</tr>
<tr>
<td>Glycine</td>
<td>33</td>
<td>12%</td>
<td>Asn (1), Ala (2), Thr (1)</td>
<td></td>
</tr>
<tr>
<td>Alanine</td>
<td>37</td>
<td>22%</td>
<td>Lys (1), Val (1), Ser (2), Thr (3), Gln (1)</td>
<td></td>
</tr>
<tr>
<td>Valine</td>
<td>30</td>
<td>13%</td>
<td>Ile (2), Tyr (1), Ala (1)</td>
<td></td>
</tr>
<tr>
<td>Isoleucine</td>
<td>13</td>
<td>38%</td>
<td>Leu (1), Thr (2), Val (2)</td>
<td></td>
</tr>
<tr>
<td>Leucine</td>
<td>15</td>
<td>29%</td>
<td>Val (1), Met (1), Tyr (1)</td>
<td></td>
</tr>
<tr>
<td>Methionine</td>
<td>5</td>
<td>20%</td>
<td>Phe (1)</td>
<td></td>
</tr>
<tr>
<td>Tyrosine</td>
<td>10</td>
<td>20%</td>
<td>Phe (1), Leu (1)</td>
<td></td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>3</td>
<td>33%</td>
<td>Asn (1)</td>
<td></td>
</tr>
<tr>
<td>Tryptophan</td>
<td>3</td>
<td>67%</td>
<td>Gly (1), Leu (1)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>275</td>
<td>30.9%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Residues shown in **boldface** type involve at least two base changes in the triplet codon; other changes can be accounted for by single base changes.

* Includes deletion.
There are segments of the sequence (present in both subtilisins) that include several hydrophobic residues in linear order. Although both subtilisins contain 13 arginine plus lysine residues, only 7 lysine and 2 arginine residues are present at the same positions in both sequences. Thus, only 5 of the 14 tryptic peptides from each of the two proteins possess the same number of residues. This was one of the reasons why it was necessary to determine independently the complete sequence of subtilisin Carlsberg. Although the study of peptide patterns from tryptic digests is very useful in instances when only a few amino acid replacements are present, the assessment of homology, when almost two-thirds of the peptides differ, is almost impossible. Indeed, the two subtilisins show closer homology by amino acid analysis than by peptide patterns.

Most of the amino acid replacements are distributed throughout the sequences of the two subtilisins. There are, however, segments of the two proteins that show long sequences with no substitutions. These regions comprise residues 64 through 75, including 2 histidines; residues 119 through 128, with 2 methionines; residues 146 through 155, with the unusual reiteration of valine and alanine residues; residues 218 through 240, containing the serine at position 221 reactive with diisopropyl fluorophosphate, as well as 2 histidine residues; and residues 261 through 270, with 3 aromatic residues in continuous sequence. Although the constancy of these segments of the sequence suggests a possible importance for conformation or enzymic function, it is premature to speculate on the significance of each of these regions.

It is noteworthy that 3 of the 5 histidine residues in both subtilisins are adjacent to proline residues: His-Pro at residues 39 and 40 and at residues 226 and 227, and Pro-His at residues 119 through 128. Clearly, these histidine residues are not likely to be present in an $\alpha$-helix or $\beta$-structure.

Since the most sensitive substrates of subtilisin are those possessing aromatic side chains (13, 14), it is likely that the substrate-binding site for such side chains includes a hydrophobic area. There are three segments of the sequence (present in both subtilisins) that include several hydrophobic residues in linear sequence: residues 147 through 150 (Val-Val-Val-Val), residues 233 through 235 (Leu-Ile-Leu), and residues 261 through 263 (Phe-Tyr-Tyr); however, these areas could just as well be important for the conformation of the protein itself. Nonetheless, the findings that the $K_m$ values for different substrates for the two subtilisins differ only in minor ways (14) suggests that the binding site or sites must be essentially similar in the two enzymes.

**Repetitions in Sequences**

A striking feature of the subtilisins is the occurrence of a number of repetitions in different segments of the sequence. These are apparent in both subtilisins and, indeed, are somewhat more obvious when both proteins are considered together (Fig. 2). We omit all repetitions that would necessitate assuming extensive deletions (or additions) of residues. The most important of such repetitions are shown in Fig. 3.

For present purposes we assume correspondence of positions if one subtilisin or the other shows an identical residue in the sequences being compared. We have included only one tetrapeptide sequence. Although we put no emphasis at present on tripeptide sequences, we should note that there are a number of repetitions of this type, e.g. Ala-Ala-Ala (residues 151 through 153, 230 through 232, and 272 through 274), Val-Lys-Val (residues 26 through 28 and 93 through 95), Gly-Asn-Ser (residues 154 through 156 and 157 through 159 in Carlsberg), and Val-Leu-Gly (residues 81 through 83 and 95 through 97 in BPN').

The sequences shown in Fig. 3A correspond in seven of the nine positions, and there is one replacement that might be interpreted as conservative (Val $\leftrightarrow$ Ala). The sequences in Fig. 3B manifest identical residues in seven positions, assuming two deletions (or additions); at least three of the remaining sites are occupied by similar residues. Eight of the 12 sites in the sequences shown in Fig. 3C are occupied by identical residues. The shorter sequences also manifest striking similarities in sequence, the most interesting being $D$, involving 2 of the histidine residues. Thus, 4 of the 5 histidines present in both subtilisins are involved in repetitive sequences.

The presence of repetitive sequences in other proteins has generally been interpreted as involving an extension of shorter peptide chains by a process of duplication of nucleotide sequences within a gene; some possible mechanisms for achieving this have been postulated.

---

**Fig. 3.** Segments of the sequences of subtilisins BPN' and Carlsberg which appear to involve repetition of similar sequences. Residues shown in **boldface** type are identical; residues shown in *italics* would appear to represent conservative replacements. The continuous sequences are BPN' or both strains; residues shown above or below these sequences represent differences in the Carlsberg strain.
been reviewed by Dixon (15). For the subtilisins, which consist of a single long peptid chain, it is reasonable to assume that at some earlier stage of evolutionary development chain extension occurred by a duplication of sequence in the DNA.

There are some features of the repetitions shown in Fig. 3 that warrant comment. First, the repetitions are not in exact linear order, but for some segments appear to be in inverse order. For example, if we consider the sequences in A, B, and D, the duplicated sequences are in the order residues 39 through 42, 67 through 75, and 82 through 94 in the NH2-terminal part of the molecule but are repeated in the inverse order: residues 238 through 241, 226 through 233, and 126 through 136. Second, some parts of the sequences appear to be twice repeated; e.g. residues 82 through 88 (Fig. 3, B and C) are repeated not only by residue 126 through 131 but also by residue 168 through 174. Similarly, residues 126 through 130 (Fig. 3E) show homology with residues 45 through 49. Thus, essentially the same pentapeptide sequence occurs four times, if we assume that residue 84 (valine) represents a late addition to the sequence.

The fact that the numbers of residues between the repeated segments of A, B, and D differ in the two parts of the molecule, as well as their occurrence in inverse order, suggests that the subtilisins evolved from simpler precursors by rather complex processes. Furthermore duplications of the sequence may have occurred several times. It is hoped that knowledge of the sequences of other homologous subtilisins will aid in extending our knowledge of the origin and evolutionary development of these enzymes and provide more detailed insight into the processes of such development.

**RELATIONSHIPS WITH OTHER PROTEINASES**

One of the major reasons for undertaking a study of the sequences of the subtilisins was to determine what relationships, if any, these enzymes have to the pancreatic proteinases. The known complete sequences of trypsin (and trypsigen) (16), (17) and chymotrypsin (and chymotrypsigen) (17), and the partially known sequences of chymotrypsin B (17, 18), indicate that all these enzymes evolved from a common ancestral precursor (16, 17, 19). The subtilisins resemble these animal enzymes in many properties, e.g. inhibition by reaction of a serine residue with diisopropyl fluorophosphate (7), involvement of histidine (13), and pH optimum. The subtilisins differ in possessing a broader specificity (14, 20). The chloromethyl ketones, inhibitors of trypsin (21) and chymotrypsin (22), do not react with the subtilisins, nor do other specific inhibitors of trypsin and chymotrypsin, such as proflavin (23), thionine (24), and Biebrich scarlet (25). In primary structure, the animal and bacterial enzymes appear to be almost totally different. Sanger and Shaw (26) noted that the tetrapeptide sequence around the reactive serine in subtilisin Novo* differs from that of the pancreatic proteinases. Unlike the pancreatic enzymes, which are rich in disulfide bridges, the subtilisins are completely devoid of cystine (or cysteine). Furthermore, the sequences of the subtilisins, considered grossly or in detail, show no greater resemblance to the pancreatic enzymes than do any group of unrelated proteins. This is particularly striking for the sequence in the region of the serine reactive with diisopropyl fluorophosphate as well as for the sequences near the 5 constant histidine residues.

It is now apparent for many proteins that are homologous in function in the vertebrates or, indeed, in several phyla, that evolutionary relationships are manifested by very similar amino acid sequences. Most extensive data are available for the cytochromes c, and it is evident that this protein has retained its function and the critical features of its sequence during the evolution of Ascomycetes, plants, and animals (10, 11). Less extensive data for many other proteins show similar retention of sequences. Thus, it was somewhat surprising to find that this does not appear to be the case for the proteinases of animal and bacterial origin possessing a reactive serine. At present, we can only suggest that the animal and bacterial enzymes may have evolved independently. If so, this would represent a possible instance of parallel evolution at the molecular level, a phenomenon well known in morphological evolution. Moreover, proteinases of other types, e.g. papain (27) or the related rennin and pepsin (28, 29), would also appear to have evolved independently, as judged from the presently available sequence information.

During the evolution of the pancreatic proteinases, a clear differentiation of specificity developed, despite the similarities in amino acid sequences, e.g. between trypsin and chymotrypsin. For the two subtilisins of known sequence, the specificity is very similar, as judged by studies on polypeptide substrates, or on acylamino acid esters or amino acid esters (14). Yet there are significant differences, the Carlsberg enzyme possessing a higher kcat for the deacylation reaction with the N-trans-cinnamoyl derivative and higher Vmax values for several substrates (14). Furthermore, the relative rates for several substrates show some differences for the two enzymes. This suggests that at least some of the 85 differences in sequence are present in parts of the molecules involved in substrate binding or near the reactive sites. Thus, some differentiation of these two enzymes has occurred, and this may be regarded as a necessary intermediate stage in evolution before a sharp distinction in specificity develops.

From the study of the electron-transferring cytochromes c (10, 11) and the oxygen-binding hemoglobins (30), it became evident that considerable variation in sequence was possible without significant change in function. Our investigations of two subtilisins has shown that this is also the situation with enzymes. Another example has recently been provided for bovine and rat ribonucleases (31). Thus, most proteinases exhibit a great tolerance for amino acid substitutions. Only radical amino acid substitutions that influence critical features of the conformation or the active sites of enzymes would not be expected to allow retention of function (12). The study of tryptophan synthetases by Yanofsky et al. (32) has provided an excellent example in an enzyme of the study of amino acid substitutions that modify or destroy function. Obviously, the two types of studies on altered proteins provide complementary information.

---

*F. S. Markland and E. L. Smith, unpublished studies.


4M. Otteeen, personal communication.
REFERENCES

Subtilisin Carlsberg: V. THE COMPLETE SEQUENCE; COMPARISON WITH SUBTILISIN BPN'; EVOLUTIONARY RELATIONSHIPS
Emil L. Smith, Robert J. DeLange, W. Howard Evans, Michael Landon and Francis S. Markland


Access the most updated version of this article at http://www.jbc.org/content/243/9/2184

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC’s e-mail alerts

This article cites 0 references, 0 of which can be accessed free at http://www.jbc.org/content/243/9/2184.full.html#ref-list-1