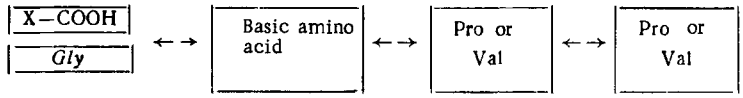
STRUCTURALLY AND FUNCTIONALLY IDENTICAL FRAGMENTS IN PEPTIDE HORMONES AND KININS

G. I. Chipens and I. M. Liepinya

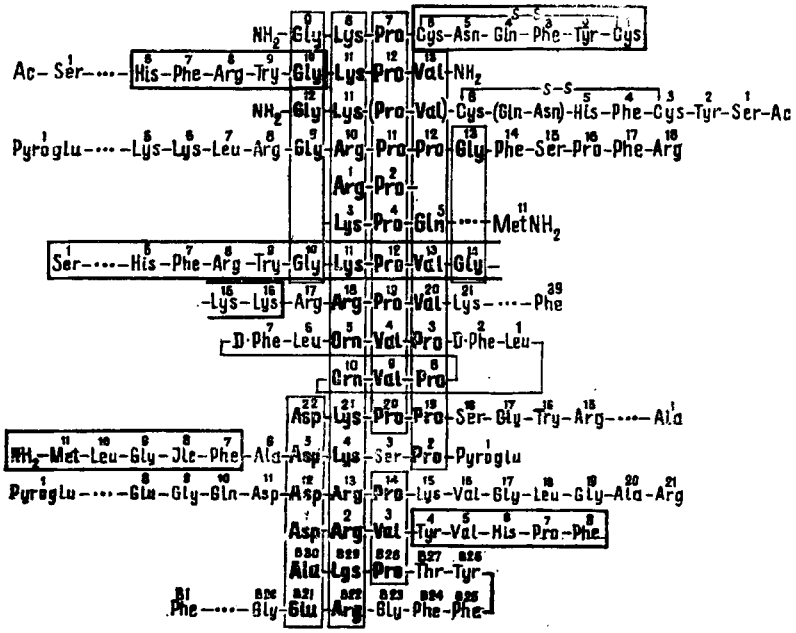
UDC 547.946.4

An analysis that we have performed of the amino acid sequences of a number of physiologically active peptides differing both in origin and also in structure and in the functions fulfilled (Scheme 1) has shown [1] that a number of peptides contain fragments including proline and valine, a basic amino acid (Arg, Lys, Orn, or Abu), and an amino acid with a free carboxy group or glycine.



(1)

We have provisionally denoted these fragments by the term "common" fragments.



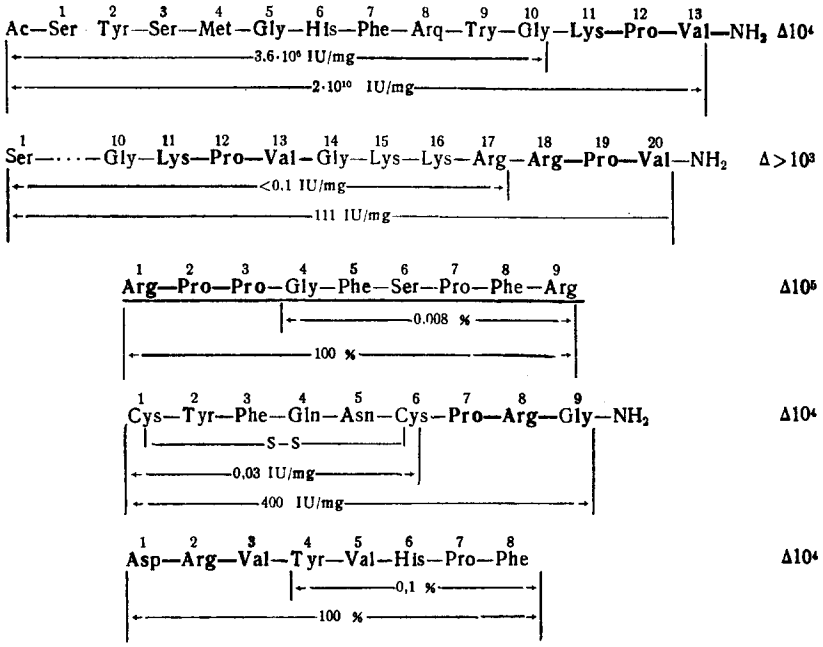
Scheme 1

Primary structures of some peptide hormones and kinins (ADH, α-MSH, CRF, wasp kinin, substance II, ACTH, gramicidin C, β-MSH, eledoisin, fibrinopeptide B, angiotensin, insulin). The common fragments (vertical columns) and the Hofmann active centers are shown in heavy type.

The great functional importance of the common fragments found is shown by the fact that their removal from or addition to a peptide molecule changes its biological activity by an average of three or four

Institute of Organic Synthesis, Academy of Sciences of the Latvian SSR. Translated from Khimiya Prirodnykh Soedinenii, No. 1, pp. 84-88, January-February, 1973. Original article submitted April 17, 1972.

orders of magnitude. Thus, for example, the addition of the common tripeptide basic amino acid–proline–valinamide potentiates the effects of the ACTH and the MSH fragments equally strongly (Scheme 2). In view of the fact that the structures of these and other common fragments are identical or extremely similar, it may be assumed that the mechanism of the potentiation of a biological effect by these fragments is the same in all cases. It is possible that the common fragments take a direct part in the formation of the secondary signal and reflect some principles or other of a universal physicochemical mechanism of the action of a number of groups of physiologically active peptides [1, 2].



Scheme 2

Biological activity of some peptides before and after the addition of a common fragment: 1) MSH (melanotropic activity); 2) ACTH (with ascorbic acid); 3) bradykinin (depressor activity); 4) ADH (pressor activity); 5) angiotensin (pressor activity); Δ represents the change in activity [1].

For further investigations, it was interesting to consider the statistical probability of finding in the structures of the peptides studied two or more monotypical amino acids and also to analyze selectively the primary structures of some proteins for the presence of common fragments in them.

Table 1 shows calculations of the probability coefficients for two "arbitrary" decapeptides and for the concrete peptides shown in Scheme 1. As can be seen, if the number of amino acids in the common fragment exceeds two, the statistical probability of finding fragments with identical structures becomes very slight (statistically it is possible to find two decapeptides with an identical sequence of three amino acids only in sets containing not less than 1067 decapeptides). The probability of finding the sequence Arg–Pro is 1067(1/20) m.

Here, m is the number of amino acids in the common fragment. Although the coefficient of probability increases with an increase in the length of the chain and with the admission of retrosequences of amino acids, nevertheless, when one takes the limited number of peptide hormones and kinins of animal organisms into account the probability of finding common fragments is extremely slight. Consequently, it can be stated that the presence of common fragments in the structures of peptide hormones is not fortuitous but reflects some general law.

The analysis of the primary structure of 14 arbitrarily selected proteins [3] of different origin and function (Table 2) showed that common fragments of type I are characteristic mainly only for low-molecular-weight peptide hormones and kinins. Among the more than 4000 sequences of tripeptides included in the structures of these proteins, only one fragment of type I (Lys-Pro-Val) was found (in the ribonuclease

TABLE 1. Probability of Finding Common Fragments in the Structures of Low-Molecular-Weight Peptides

No. of amino acids in the peptide		No. of amino acids in the common fragment	Prob. of finding any common fragment in the peptides N and M
N	M		
10	10	2	0,2025
10	10	3	0,0080
9(ADH)	13(MSH)	3	0,0096
9(ADH)	39(ACTH)	3	0,0324
12(CRF)	13(MSH)	4	0,0005
12(CRF)	39(ACTH)	4	0,0020
18 (Wasp kinin)	39(ACTH)	5*	0,0001

* The calculation was performed by considering Arg/Lys and Pro/Val as functionally identical amino acids.

TABLE 2. Characteristics of the Proteins Used for Analyzing Proline- and Valine-Containing Sequences [3]

Protein and its source	Tot. No. of amino acid residues	No. of residues of	
		proline	valine
Cytochrome C (human)	104	4	3
Rubredoxin (Micrococcus aerogenes)	53	4	4
Azurin (Pseudomonas fluorescens)	128	4	10
Ferredoxin (Micrococcus aerogenes)	56	5	4
Gameritin (Golfingia Gouldi)	114	4	4
α-Hemoglobin (human)	141	7	13
Bence-Jones protein (mouse, 41)	214	9	9
Trypsinogen (ox)	229	8	18
Subtilisin (Bac. subtilis BPN)	275	14	30
α-Tryptophan synthetase (Escherichia coli)	268	19	17
Lactalbumin (ox)	123	2	6
Lysozyme (chick)	129	2	6
Ribonuclease (ox)	124	4	9
Tobacco mosaic virus	158	8	14

molecule). It can be seen from a consideration of Scheme 1 and Table 2 that in the structures of the low-molecular-weight peptides mentioned proline is found on an average two and a half times more frequently (as a percentage) than in the high-molecular-weight proteins. Furthermore, in contrast to the peptide hormones and kinins, in the proteins arginine or lysine is very rarely adjacent to proline. In the low-molecular-weight peptides, proline is apparently necessary to determine their specific "biologically active" spatial conformation of the whole molecule and, particularly, of an adjacent basic amino acid. The conformation of proteins is determined mainly by the sum of the weak intermolecular interactions of numerous amino-acid residues.

The probability P has been defined as the ratio of the number of favorable outcomes to the number of all possible outcomes from the formula [3],

$$P = 1/20^m (n_1 - m + 1)(n_2 - m + 1) ,$$

where m is the number of amino acids in the whole fragment; n_1 is the number of amino acids in peptide N; n_2 is the number of amino acids in peptide M; and 20 is the number of natural amino acids. In calculations of the probability factors, we took the following into account: all peptides are formed from only 20 natural amino acids; peptides differ both in the number of types of amino acids and in the number of amino acids and their arrangement; one and the same amino acids in a chain (including common fragments) may be repeated; common fragments may be located at any position of the peptide chain and the direction of acylation of the chain may be the same or the opposite. The results of the calculations are given in Table 1.

For analyzing the primary structures of proteins (see Table 2) from their amino acid sequences [4] all the tripeptides containing proline and valine were written down and the amino acids adjacent to proline and to valine were recorded. Thus, for example, the tripeptides of cytochrome C $Asn^{70}-Pro^{71}-Lys^{72}$ and $Asp^2-Val^3-Glu^4$ were recorded in the cells Pro-Asn, Pro-Lys, Val-Asp, Val-Glu, etc. The frequency of

amino acids forming the closest neighbors of proline and valine in the structures of some proteins (see the list in Table 2) are shown below:

Neighboring amino acids	Ala	Arg	Asp	Asn	Cys	Glu	Gln	Gly	His	Ile
Central amino acids	18	1	14	13	5	4	2	18	8	12
	26	9	25	11	10	13	12	25	7	13
Neighboring amino acids	Leu	Lys	Met	Phe	Pro	Ser	Thr	Try	Tyr	Val
Central amino acids	11	7	2	9	6	14	12	1	11	13
	19	23	3	13	15	25	20	6	8	19

SUMMARY

The probability coefficients for finding fragments including identical amino acid sequences in the structures of peptide hormones and kinins have been calculated.

It has been shown that common fragments of the type of proline or valine–basic amino acid–acidic amino acid or glycine are characteristic only for structures of individual groups of peptide hormones and kinins, being rarely found in the primary structures of proteins.

LITERATURE CITED

1. G. I. Chipens, in: The Chemistry and Biology of Peptides [in Russian], Riga (1971), p. 23.
2. G. I. Chipens and O. S. Pansuevich, in: The Chemistry and Biology of Peptides [in Russian], Riga (1971), p. 5.
3. F. Mosteller, P. Rurke, and J. Tomas, Probability [Russian translation], Moscow (1969).
4. M. O. Dayhoff and R. V. Eck, Atlas of Protein Sequence and Structure 1967-1968, Silver Spring, Maryland (1969).