

多肽激素

盛树力 主编

的当代理论和应用

Contemporary
Theory and
Application of Peptide
Hormones

科学技术文献出版社

Q57
SSL

多肽激素的当代理论和应用

主编 盛树力

编者 (以姓氏笔划为序)

马金城 王 蓉 王霄虹

刘 卫 齐红伟 李 林

郭育红 赵咏梅 梅学文

姬志娟 盛树力 童文凯

温绍君



A0291460

科学技术文献出版社

(京)新登字 130 号

责任编辑/ 李 洁

策划编辑/ 陈玉珠

责任校对/ 陈悦楼

封面设计/ 雪 梅

图书在版编目(CIP)数据

多肽激素的当代理论和应用/盛树力主编. - 北京:科学技术文献出版社,1998.9

ISBN 7-5023-3085-2

I. 多… II. 盛… III. 多肽激素-基本知识 IV. Q57

中国版本图书馆 CIP 数据核字(98)第 16343 号

出 版 者/ 科学技术文献出版社

地 址/ 北京市复兴路 15 号(中央电视台西侧)/100038

发 行 者/ 新华书店北京发行所

印 刷 者/ 北京金特印刷厂

版(印)次/ 1998 年 9 月第 1 版,1998 年 9 月第 1 次印刷

开 本/ 787×1092 16 开

字 数/ 640 千

印 张/ 25

印 数/ 1—1500 册

定 价/ 50.00 元

© 版权所有 违法必究

(购买本社图书,凡字迹不清、缺页、倒页、脱页者本社发行部负责调换)

发行部电话/(010)68514035 总编室电话/(010)68515544-2935

社长室电话/(010)68515037

前 言

多肽是涉及生物体内各种细胞功能的生物活性物质。自从能用固相法人工合成多肽以来的 40 年间,伴随分子生物学、生物化学技术的飞速发展,多肽研究取得了惊人的、划时代的进展。生物体内已发现的多肽(100 个氨基酸残基以下组成)已达几百种,是机体完成各种复杂的生理活动必不可少的参与者。所有细胞均能合成多肽类物质,也受多肽所调节,所以多肽是生命科学研究中的热点之一。多肽研究不仅涉及生物和化学合成、生物活性、受体及信号传导,而且展示了临床应用的巨大潜力,已成为集多学科的一门专业。我国多肽研究有过辉煌过去,也有今天的成就,如果多肽学者和医学界、生物学界甚至包括有远见的企业家形成完美的结合,明天的多肽研究将以高学术水平走向世界和以高技术水平走向市场。我们编写这本书的目的除为我校研究生提供正规教材外,更重要的目的还在于使多肽研究工作者和医学、生物学工作者的知识互相补充,促进相互了解,达到基础理论和临床应用的共同发展。但由于水平的限制,我们的目的很难完全达到,所以本书只是抛砖引玉。如能引起国内专家和科研支持部门的一点注意就足以满足我们编写本书的愿望。

本书由首都医科大学宣武医院、北京市老年病医疗研究中心从事多肽研究的老、中、青科研工作者编写。我相信本书一定存在许多缺点和我们没有发现的错误,望同道不吝指正。

本书的编写得到了我院领导的支持,赵志伟同志负责本书的打印和整理文稿的工作,俞乃昌教授对英文图表进行了翻译,在此表示感谢。

盛树力

首都医科大学宣武医院
北京脑老化研究实验室

1998.1

目 录

第一章 多肽的基本知识和研究范畴	(1)
一、多肽的化学性质和稳定性	(1)
二、多肽的 mRNA 前体的剪接和译后加工	(5)
三、多肽的种间差异和同功能多肽的分子间差异	(7)
四、多肽构效关系	(8)
五、合成多肽片段测定体内蛋白质抗体	(10)
六、测定已知序列的自体蛋白质	(11)
七、多肽测定	(11)
八、新多肽的发现	(15)
九、多肽的分泌	(16)
十、神经、体液和免疫之间的联系	(17)
第二章 多肽激素的生物合成	(20)
一、DNA 模板的基本概念	(20)
二、RNA 转录	(21)
三、蛋白质与多肽类激素的生物合成	(25)
四、蛋白质与多肽激素生物合成的调节	(31)
第三章 多肽的化学合成和纯化	(33)
一、多肽的结构与分类	(33)
二、多肽合成的基本原理及肽键的生成	(34)
三、固相肽合成	(35)
四、多肽研究中的进展	(45)
五、多肽的高效液相色谱分离与制备	(46)
第四章 肽类激素作用的机制	(55)
一、膜受体	(55)
二、受体信号整合和控制	(59)
三、受体效应系统	(60)
四、胞浆内第二信使	(65)
第五章 下丘脑垂体激素概论	(70)
一、定义	(70)
二、合成和生物学特性	(70)
三、下丘脑中其他的活性物质	(73)
四、垂体激素细胞及其相互作用	(74)
五、垂体细胞的个体发育	(75)
六、下丘脑垂体激素与临床	(75)
第六章 生长激素释放激素和生长激素释放抑制因子	(78)
一、GHRH 基因的结构	(78)

二、GHRH 基因表达和分泌的调节	(79)
三、GHRH 的功能	(81)
四、GHRH 和 GHRP 的应用	(84)
五、生长激素释放抑制因子(生长抑素)	(86)
第七章 人生长激素	(96)
一、人生长激素基因族	(96)
二、人生长激素的基因结构	(97)
三、GHF-1	(99)
四、GH 基因转录的调节控制	(103)
五、hGH 基因的表达	(108)
六、胎盘中 GH 基因家族的表达	(109)
七、hGH 基因的缺失	(110)
八、hCS 和 hGH-V 基因缺失	(111)
九、GH 受体	(112)
十、GH 结合蛋白和 GH 受体结合蛋白	(114)
十一、GH 分泌的调节和临床应用	(115)
第八章 泌乳素	(121)
一、泌乳素结构	(121)
二、泌乳素基因表达的组织特异性调节	(122)
三、泌乳素的产生部位	(123)
四、泌乳素分泌的调节	(124)
五、泌乳素的作用	(125)
六、泌乳素受体	(126)
七、泌乳素测定的临床意义	(129)
第九章 促甲状腺素释放激素和甲状腺刺激激素	(130)
一、促甲状腺素释放激素	(130)
二、甲状腺刺激激素	(138)
第十章 促肾上腺皮质激素释放因子和促肾上腺皮质激素	(148)
一、促肾上腺皮质激素释放因子	(148)
二、促肾上腺皮质激素	(161)
第十一章 促性腺激素释放激素和促性腺激素	(168)
一、促性腺激素释放激素	(168)
二、促性腺激素	(173)
三、促性腺激素释放激素的类似物和临床应用	(185)
第十二章 垂体后叶激素	(190)
一、神经垂体激素的合成部位	(190)
二、基因结构及表达	(190)
三、化学结构及前体	(191)
四、受体	(192)
五、AVP 类似物	(194)

六、生物合成、分布与代谢	(195)
七、神经分泌细胞与神经肽的释放	(196)
八、AVP的释放调节	(196)
九、AVP的作用	(200)
十、OXT的释放调节	(201)
十一、OXT的作用	(202)
十二、神经垂体激素的行为效应	(202)
第十三章 胰岛素	(204)
一、胰岛素的结构和生物活性	(204)
二、胰岛素基因的表达和调控	(206)
三、胰岛素的生物合成和分泌	(210)
四、胰岛素的代谢	(216)
五、胰岛素受体和受体后信号传导	(216)
六、胰岛素的作用	(223)
七、胰岛素抵抗和Ⅱ型糖尿病	(226)
八、胰岛素测定的临床意义和人用胰岛素制剂	(233)
第十四章 胰升糖素及相关肽	(236)
一、分子生物学	(236)
二、胰升糖素	(239)
三、胰升糖素样肽-1	(241)
四、中枢神经系统对血糖浓度的调节	(242)
五、GLP-1受体及信息传递	(244)
六、glicentin、OM和KA8	(246)
第十五章 胰岛素样生长因子	(249)
一、胰岛素样生长因子	(249)
二、IGF受体	(257)
三、IGF结合蛋白	(261)
四、IGF体系的生理功能	(266)
五、IGF-I临床应用前景	(272)
第十六章 降钙素基因相关肽和肾上腺髓质素	(275)
一、降钙素基因相关肽	(275)
二、肾上腺髓质素	(284)
第十七章 胰淀粉样多肽	(298)
一、分子生物学	(298)
二、分子结构	(300)
三、分布	(302)
四、合成与分泌	(302)
五、生物学作用	(303)
六、在糖尿病中的作用	(306)
第十八章 利钠激素	(309)

一、分子生物学	(309)
二、利钠多肽的译后加工和分布	(311)
三、利钠激素受体	(314)
四、利钠多肽的生物活性	(316)
五、利钠激素分泌的调节	(318)
六、利钠激素的病生理意义	(319)
第十九章 钙调节激素	(322)
一、甲状旁腺激素	(322)
二、甲状旁腺激素相关蛋白	(328)
三、甲状旁腺激素和甲状旁腺激素相关蛋白的受体	(332)
四、降钙素	(335)
第二十章 胸腺激素	(341)
一、胸腺器官和细胞	(342)
二、胸腺激素的生物化学	(343)
三、胸腺激素的生物学作用	(349)
四、胸腺激素的治疗学作用	(355)
五、结束语	(363)
第二十一章 肾素-血管紧张素体系	(365)
一、RAS 代谢链的分子生物学	(365)
二、循环 RAS 及其作用	(370)
三、局部 RAS 及其作用	(371)
四、循环 RAS 与局部 RAS	(375)
五、RAS 抑制剂	(376)
第二十二章 内皮素	(379)
一、概述	(379)
二、ET 的生物学作用	(379)
三、ET 与心脑血管病	(382)
四、ET 作用的受体机制	(385)
五、ET 的 cDNA 及其基因结构	(389)
六、ET 受体拮抗剂的研究	(390)

第一章 多肽的基本知识和研究范畴

多肽是一类生物活性物质,是生命活动必不可少的参与者,它涉及激素、神经、细胞生长和生殖等各个领域,其重要性在于调节体内各系统、器官和细胞的功能活动,因此研究多肽及其受体的生物合成、分布和功能就成为近二三十年来科学家注意的中心之一。

多肽和蛋白质只是肽链长短之别,二者没有严格的区分。通常由少于 100 个氨基酸组成的称为多肽,而由 100 个以上氨基酸组成的为蛋白质,它们的主链都称多肽链。多肽与蛋白质比较,功能有些不同。首先,它是体现信息的信使,以引起各种各样不同时效的正性或负性生理活动和生化反应的调节;第二,活性高, 10^{-7} mol/L 或以下就可以发挥活性;第三,分子小,易于结构改造,易于化学合成;第四,可以用多肽片段为工具,研究蛋白质的性质。由于这些特点,多肽研究正逐渐发展为独立的专业,它包括多肽的分子生物学、化学合成、生物活性、病理意义和多肽药物以及特异受体与信息传递等各个方面。多肽研究是一门药物化学和生物学的交界学科。

一、多肽的化学性质和稳定性

生物合成过程中,在基因遗传密码指导下,按一定序列将氨基酸通过肽键连接形成肽链(一个氨基酸的 α 羧基和另一个氨基酸的 α 氨基脱水,在肽链中氨基酸称为氨基酸残基)。多肽结构的书写方法是含 α 氨基的氨基酸一端写在左边,称氨基端或 N 端,含 α 羧基的氨基酸则写在右边,称羧基端或 C 末端,少数多肽链中还有二硫键。

氨基酸共有 20 种,可分为中性、酸性和碱性三类(表 1-1)。除 N 端和 C 端具有游离的氨基和羧基外(常酰胺化),肽链中侧链也有氨基和羧基,多肽分子中氨基数超过羧基数即为碱性多肽,反之为酸性多肽。中性氨基酸中含疏水性 R 基团的氨基酸称疏水性或非极性氨基酸,包括脂肪族 5 种:丙氨酸(A)、缬氨酸(V)、亮氨酸(L)、异亮氨酸(I)、蛋(甲硫)氨酸(M),芳香族一种:苯丙氨酸(F),杂环族 2 种:色氨酸(W)、脯氨酸(P),共 8 种。R 基团有极性但不能解离的为不解离的极性氨基酸,包括含羟基的丝氨酸(S)、酪氨酸(T)、苏氨酸(Y),含巯基的半胱氨酸(C),酰胺类氨基酸谷氨酰胺(Q)、天冬酰胺(N)和甘氨酸(G),共 7 种。酸性氨基酸有 2 种,为谷氨酸(E)和天冬氨酸(D)。碱性氨基酸有 3 种,为组氨酸(H)、赖氨酸(K)和精氨酸(R)。

任何蛋白质和多肽的主链的原子是相同的,而侧链则互不相同,侧链的重要特性是疏水性或亲水性,疏水性是指氨基酸侧链避开与水接触的程度,而亲水性恰好相反,所以实际上亲、疏水性是侧链在极性和非极性两相中分散的倾向性。肽链一级结构或氨基酸序列是决定肽链空间构象的决定因素,而氨基酸的侧链避开与水接触的程度是稳定多肽结构的一个重要因素。

氨基酸在溶液中的亲水性见表 1-2。但在膜蛋白的转膜区被认为是 α 螺旋,却有大量异亮氨酸和缬氨酸,可能表明类脂双层中螺旋形成与水根本不同,不带电荷的氨基酸在膜环境中的螺旋倾向为异亮 > 亮 > 缬 > 蛋 > 苯丙 > 丙 > 谷氨酰胺 > 苏 > 丝 > 天冬酰胺 > 甘 > 脯,这提示 α 螺旋倾向取决于侧链疏水性和多肽片段残基的水化程度。

表 1-1 氨基酸的结构和分类

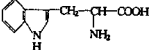
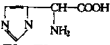
分类	名 称	结 构 式	相对分子质量	等电点
中性:	甘氨酸 glycine(甘, Gly, G)	$\begin{array}{c} \text{CH}_2\text{---COOH} \\ \\ \text{NH}_2 \end{array}$	75.07	5.97
	丙氨酸 alanine(丙, Ala, A)	$\begin{array}{c} \text{CH}_3\text{---CH---COOH} \\ \\ \text{NH}_2 \end{array}$	89.09	6.00
	丝氨酸 serine(丝, Ser, S)	$\begin{array}{c} \text{CH}_2\text{---CH---COOH} \\ \quad \\ \text{OH} \quad \text{NH}_2 \end{array}$	105.09	5.68
	半胱氨酸 cysteine(半胱, Cys, C)	$\begin{array}{c} \text{CH}_2\text{---CH---COOH} \\ \quad \\ \text{SH} \quad \text{NH}_2 \end{array}$	121.16	5.07
	蛋氨酸 methionine(蛋, Met, M)	$\begin{array}{c} \text{CH}_3\text{S---CH}_2\text{---CH}_2\text{---CH---COOH} \\ \\ \text{NH}_2 \end{array}$	149.21	5.74
	苏氨酸 threonine(苏, Thr, T)	$\begin{array}{c} \text{CH}_3\text{---CH---CH---COOH} \\ \quad \\ \text{OH} \quad \text{NH}_2 \end{array}$	119.21	6.53
	缬氨酸 valine(缬, Val, V)	$\begin{array}{c} \text{CH}_3\text{---CH---CH---COOH} \\ \quad \\ \text{CH}_3 \quad \text{NH}_2 \end{array}$	117.15	5.96
	亮氨酸 leucine(亮, Leu, L)	$\begin{array}{c} \text{CH}_3\text{---CH---CH}_2\text{---CH---COOH} \\ \quad \\ \text{CH}_3 \quad \text{NH}_2 \end{array}$	131.17	5.98
	异亮氨酸 isoleucine(异亮, Ile, I)	$\begin{array}{c} \text{CH}_3\text{---CH}_2\text{---CH---CH---COOH} \\ \quad \\ \text{CH}_3 \quad \text{NH}_2 \end{array}$	131.17	6.02
	苯丙氨酸 phenylalanine(苯丙, Phe, F)	$\begin{array}{c} \text{C}_6\text{H}_5\text{---CH}_2\text{---CH---COOH} \\ \\ \text{NH}_2 \end{array}$	165.19	5.48
	酪氨酸 tyrosine(酪, Tyr, Y)	$\begin{array}{c} \text{HO---C}_6\text{H}_4\text{---CH}_2\text{---CH---COOH} \\ \\ \text{NH}_2 \end{array}$	181.19	5.66
	脯氨酸 proline(脯, Pro, P)		115.13	6.30
	色氨酸 tryptophan(色, Trp, W)		204.22	5.89
酸性:	天冬酰胺 asparagine(天-NH ₂ , Asn, N)	$\begin{array}{c} \text{CO---CH}_2\text{---CH---COOH} \\ \quad \\ \text{NH}_2 \quad \text{NH}_2 \end{array}$	132	5.41
	谷氨酰胺 glutamine(谷-NH ₂ , Gln, Q)	$\begin{array}{c} \text{CO---CH}_2\text{---CH}_2\text{---CH---COOH} \\ \quad \\ \text{NH}_2 \quad \text{NH}_2 \end{array}$	132	5.65
	谷氨酸 glutamic acid(谷, Glu, E)	$\begin{array}{c} \text{HOOC---CH}_2\text{---CH}_2\text{---CH---COOH} \\ \\ \text{NH}_2 \end{array}$	181.19	3.32
	天冬氨酸 aspartic acid(天, Asp, D)	$\begin{array}{c} \text{HOOC---CH}_2\text{---CH---COOH} \\ \\ \text{NH}_2 \end{array}$	181.19	2.77
碱性:	组氨酸 histidine(组, His, H)		181.19	7.59
	赖氨酸 lysine(赖, Lys, K)	$\begin{array}{c} \text{CH}_2\text{---CH}_2\text{---CH}_2\text{---CH}_2\text{---CH---COOH} \\ \quad \\ \text{NH}_2 \quad \text{NH}_2 \end{array}$	181.19	9.14
	精氨酸 arginine(精, Arg, R)	$\begin{array}{c} \text{NH}_2\text{---C---NH---(CH}_2\text{)}_3\text{---CH---COOH} \\ \quad \\ \text{NH} \quad \text{NH}_2 \end{array}$	181.19	10.76

表 1-2 氨基酸的亲水性和疏水性尺度

氨基酸		水化指数	亲水性	疏水性
精氨酸	R	-4.50	3.00	-2.53
赖氨酸	K	-3.90	3.00	-1.50
天冬氨酸	D	-3.50	3.00	-0.90
谷氨酰胺	Q	-3.50	0.20	-0.85
天冬酰胺	N	-3.50	2.00	-0.78
谷氨酸	E	-3.50	3.00	-0.74
组氨酸	H	-3.20	-0.50	-0.40
丝氨酸	S	-0.80	0.30	-0.18
苏氨酸	T	-0.70	0.40	-0.05
脯氨酸	P	-1.60	0.00	0.12
酪氨酸	Y	-1.30	-2.30	0.26
半胱氨酸	C	2.50	-1.00	0.29
甘氨酸	G	2.50	-1.00	0.29
丙氨酸	A	1.80	-0.50	0.62
蛋氨酸	M	1.90	-1.30	0.64
色氨酸	W	-0.90	-3.40	0.81
亮氨酸	L	3.80	-1.80	1.60
缬氨酸	V	4.20	-1.50	1.08
苯丙氨酸	F	2.80	-2.50	1.19
异亮氨酸	I	4.50	-1.80	1.38

各种蛋白质和多肽在体内有不同的半生期,这就涉及肽链的稳定性。这种稳定性涉及肽链的一级结构,特别是N端的氨基酸种类,即所谓N端规则。N端中蛋氨酸、甘氨酸、丙氨酸、丝氨酸、苏氨酸和缬氨酸使肽链稳定,而赖氨酸、天冬氨酸、苏氨酸、精氨酸和脯氨酸使肽链不稳定。但肽链内部的序列也十分重要,不稳定的PEST序列(富含脯、谷、丝和苏)出现在短半生期的蛋白质中。半生期长的蛋白质中天冬酰胺、赖氨酸和甘氨酸含量较高,不稳定蛋白质中,蛋氨酸、谷氨酰胺、脯氨酸、谷氨酸和丝氨酸出现频率较高。实际上,肽链是否稳定取决于肽链的稳定性,如果某个肽键易受酶水解,则肽链不稳定。多肽一般只具有二级结构,即 α 螺旋、 β 折叠和 β 转角,小肽可能只是线形结构,而有些肽具有三级结构,如胰岛素。各种氨基酸残基出现在二级结构的3种类型的频率是不同的,即使在长链多肽中也不总是具有3种类型的二级结构,这说明氨基酸形成3种二级结构的倾向性是不同的,谷氨酸主要分布在 α 螺旋,丙氨酸和亮氨酸等在稳定的螺旋结构起重要作用,而带 α 侧链的异亮氨酸、缬氨酸和苏氨酸在 β 折叠中出现较多,天冬氨酸和甘氨酸在 β 转角出现较多,但脯氨酸常出现在 β 转角,而不出现在 α 螺旋。

球蛋白表面的 α 螺旋,而向溶剂的多肽片段比而向蛋白质内部的一面极性强得多,而跨膜区中的 α 螺旋,疏水性更强的一面渗入到膜脂质中。半生期短的多肽,如将这类氨基酸置换,就可增加肽链的稳定性,但只有那些能提高多肽稳定性而不影响其功能的置换才是有用的。有些多肽在血循环中还能与特异的结合蛋白结合而免遭酶水解从而延长半生期。

谈到多肽的功能,可以说任何多肽都是多功能的活性物质,这取决于多肽产生的部位和受体分布的部位,即同一多肽对不同靶细胞具有不同的功能,而同一生理功能绝大部分是由多种生物活性物质,包括多肽共同完成。在一些肽链的多肽序列中,可能包含具有不同生物活性的

不同肽段,即结构和活性可能析离,也有许多肽降解成不同的片段与受体结合发挥作用。如在 AVP 中 AVP4-8 有增强记忆的作用。

多肽的命名,主要根据以下几个方面:

(1) 主要的生物活性:如下丘脑和垂体分泌的催产素、抗利尿激素、生长激素、促肾上腺皮质激素等。

(2) 最初发现的部位:如内皮素、胰岛素、肾上腺髓质素等。

(3) 化学结构:如丙甘素、谷胱甘肽等。

(4) 二者结合:如赖氨酸缓激肽(胰激肽)、心钠素(心房利钠激素)等。

各种命名都有优缺点,这是显而易见的,很难既考虑到功能,又顾及产生的部位,一般都使用发现者所给予的名称。

粗略计算多肽的相对分子质量可将氨基酸数乘以 110 即可。20 种氨基酸的平均相对分子质量为 138,但多肽分子中较小的氨基酸占优势,因此平均相对分子质量接近 128,形成肽链时脱水则为 110。对已知大概相对分子质量未知序列的多肽,除以 110 则为氨基酸数的近似值。

在文献中多肽序列的表示见表 1-3,从表中可发现它们的规律。多肽序列可用氨基酸的单个英文字母或 3 个字母缩写自 N 端至 C 端写出,但通常在多肽名称后标明从第几位到第几位的数即可。如肽链没有任何变化只用名称即可。

表 1-3 多肽序列表示法

表 示 法	举 例
DRVYIHPF	全序列(AT-Ⅱ)
AT-Ⅱ 1-8	AT-Ⅱ 8 肽,无氨基酸置换
GLP-1(7-36)NH ₂	GLP-1 缺 N 端 6 个氨基酸,C 端酰胺化
Tyr-ACTH4-9	4 位氨基酸前加酪氨酸
[Asn1, Val5]-AT-Ⅱ	AT-Ⅱ 分子 1,5 位被天冬酰胺和缬氨酸取代
des-asp-AT-Ⅱ	AT-Ⅱ 1 位去天冬氨酸
desamino-Cys(去氨基-Cys)	3-mercaptopropionyl(3-巯基丙酰基)
Lys-(Ala3)-bradykinin(缓激肽)	0 位赖氨酸 3 位丙氨酸
N-Ac	N 端乙酰化
biotinyl-	N 端生物素化
[D-Ala]-LHRH	6 位 D-丙氨酸
Sar-	sarcosine(肌氨酸)
P-cl	P-chloro(氯化)
D-Cpr	D-cyclopropyl-(环丙基-)
Nle	正亮氨酸 Norleucine
D-Nal(2)	3-(2-萘基)-D-丙氨酸
Me+(O)-	蛋氨酸 sulfoxide
-M+(O)-Ol	
Sta	statine(3s,4s)-4-amino-3-hydroxy-6-methylheptanoic acid
3,5Di-Tyr	3,5-二碘酪氨酸
N-Me	N-甲基
N-t-BOC-	N-t-叔丁氧羰基
PGlu or Pyr	焦谷氨酸

二、多肽的 mRNA 前体的剪接和译后加工

1. 多肽 mRNA 剪接和翻译

多肽链是基因表达的产物。基因表达时先由 DNA 聚合酶 II 作用转录为含有外显子和内含子的前体 mRNA, 然后剪去内含子, 外显子拼接为成熟的 mRNA。剪接的位点十分严格, 以保证多肽前体具有恒定的氨基酸序列。绝大多数多肽是一个基因表达一种多肽前体, 但是少数多肽基因转录后, mRNA 前体在剪切过程中, 由于细胞中特异因子的控制, 外显子拼接位点的移位而产生不同的成熟的 mRNA, 从而产生不完全相同的多肽前体。如降钙素/降钙素基因相关肽(CGRP), 蛋白基因由 6 个外显子、5 个内含子组成。甲状腺 C 细胞的成熟 mRNA 由外显子 1~4 组成, 前体翻译后加工产生降钙素, 而在神经组织表达 CGRP 前体时的 mRNA 由外显子 1~3、5、6 组成, 二者前体的 N 端因均有外显子 1~3 表达, 所以其序列是相同的。胰岛素样生长因子 IGF-1 mRNA 前体含 5 个外显子和 4 个内含子, 经过不同的拼接, 产生 IGF-1A(外显子 1~3、5)和 1B(外显子 1~4)二种成熟 mRNA, 翻译后分别产生 153 个氨基酸的 IGF-1A 和 195 个氨基酸的 1B 前体, 所以二者的 N 端 134 个氨基酸相同, 49~118 序列为 IGF-1。二者前体的 C 端就产生不同的多肽。人生长激素 90% 由 191 个氨基酸组成, 但 mRNA 拼接时约有 5% “移位”拼接而缺失表达 32~46 位肽段的 mRNA 相应序列, 从而翻译出 176 个氨基酸(相对分子质量 20000Da)的生长激素。又如速激肽是一类具有共同 C 末端的 Phe-X-Gly-Leu-MetNH₂ 结构(X 为疏水性或芳香族氨基酸的 20 多种多肽的家族), 它们来自 2 个基因 PPT-A 和 PPT-B。PPT-A 有 α 和 β 2 种 mRNA, α 由 2~5、7 外显子拼接, β 由外显子 2~7 组成。外显子 3 翻译 P 物质, 外显子 6 翻译 K 物质, 而 mRNA 的 4、5、6 外显子的产物是 NPK。NPK C 端 10 肽即为 SK。这些事实说明不同的外显子拼接产生不同的成熟 mRNA 序列, 翻译后产生不同的多肽, 即一个基因可以产生多种活性相关不大的多肽。

新合成的多肽链称前肽原(prepropeptide), 前肽原经翻译后加工才能产生有生物活性的多肽。信号肽被信号肽酶切除后成为肽原(propeptide), 肽原再经各种酶的作用裂解出有活性的多肽和一些功能尚不明确的多肽。

信号肽通常由 15~30 个氨基酸残基组成, 但 IGF-1 的信号肽为 48 肽是个例外。信号肽分三区, N 端带正电荷的 N 区(1~5 个氨基酸), 中央疏水的 M 区(7~15 个氨基酸)和极性的 C 端 C 区(3~6 个氨基酸), C 区末端是信号肽酶的识别位点(如丙氨酸、甘氨酸)。信号肽的作用是引导刚合成的多肽链插进并通过内质网膜上的通道, 这个过程除少数蛋白质外与翻译是偶联的。

多肽原(或称前体)常包含 2 个或 2 个以上的多肽。少数前肽原包含多个重复的活性多肽序列, 如促甲状腺激素释放激素(TRH)前体中有 6 个 TRH 直接前体 TRH-GLY。

活性多肽的前体通常由几十个到 1~200 个氨基酸组成, 加工后生成 2 个或 2 个以上的小肽。前体肽链长短与活性多肽长短无关。但长肽如一些垂体前叶激素在切去信号肽之后即为活性多肽, 所以它们只有前激素(prohormone)而没有激素原(prohormone)。有些小肽的前体却来自大的蛋白质, 如血管紧张素-I(AT-I)(9 肽)来自 452 个氨基酸的前体, 缓激肽来自相对分子质量大的激肽原。

前体通常没有或几乎没有生物活性, 胰岛素原的生物活性只有胰岛素的 1/10, 所以前体

进一步加工是十分必需的,而前体的加工方式具有多样性。

2. 多肽翻译后加工

(1) 前体经蛋白酶或肽酶作用产生活性多肽:这是最常见的加工方式,但这种加工方式又常不完全,以致活性多肽可能存在 N 端或 C 端的肽链延长,出现不同分子形式的多肽,肽链中双碱性氨基酸(R-R, K-R, K-K)是胰蛋白酶样水解酶如内肽酶(endopeptidase)的酶切位点,但切开后肽 C 端尚存的一个碱性氨基酸需经碱性氨基酸特异性氨基肽酶和羧基肽酶切掉。现已发现成对的碱性氨基酸特异的内肽酶可切去成对的碱性氨基酸,在酵母为 KEX2,在哺乳动物为 PC2 和 PC3,但有些多肽中成对碱性氨基酸不被水解,如甲状旁腺素(PTH)前体、胰高血糖素前体。可能这些前体的立体构型使酶不易接近这类肽链。但其他的氨基酸也可作为酶切位点,如 AT 原在肾素作用下在亮氨酸-亮氨酸(Leu-Leu)处裂解生成 AT-I。

有些肽在二次或多次加工后,才出现活性更大、与受体有高亲和力的活性多肽。如截断型 GLP-1, TRH 前体先酶解为 TRH 直接前体 TRH-GLY,再成为 TRH。如血管紧张素原需二次酶切才产生出 AT-II。

(2) 二硫键形成:因 mRNA 没有胱氨酸密码,肽链合成后 2 个半胱氨酸由二硫异构酶催化形成二硫键,二硫键都在链内形成。胰岛素前体直链多肽链内 3 个二硫键形成后,切去 C 肽即成为有 2 个链间、一个链内二硫键的胰岛素。

(3) C 端酰胺化:邻接 C 端的甘氨酸是 C 端最后一个氨基酸的酰胺基供体,多肽 C 端酰胺化常使生物活性增强。

(4) 残基侧链的修饰:①环化:N 端谷氨酸常自发环化为焦谷氨酸。②磺酸化(SO_3H):如缩胆囊素(CCK)、胃泌素中酪氨酸的磺酸化。③磷酸化:ACTH 可能被磷酸化。④羟化:mRNA 无羟脯氨酸、羟赖氨酸等的遗传密码,肽链中出现这些氨基酸都是翻译后羟化的结果。⑤ N 端乙酰化: β 内啡肽可因乙酰化丢失活性,而 α 促黑激素(α -MSH)活性增强。⑥糖化:不少分泌性蛋白质常被糖化形成糖化蛋白,但糖化位点的氨基酸序列有一定模式,如 N-糖基化位点是 N-X-S 或 T 的三联结合,第一个是天冬酰胺,第二个是除脯氨酸、门冬氨酸以外任一氨基酸,但第三个一定是丝氨酸或苏氨酸。并不是具备三联结构就能被糖化,它还受到蛋白质立体构型的制约。有些糖蛋白通过丝氨酸或苏氨酸的 OH 基与 N-乙酰半乳糖胺(N-acetyl-galactosamine)结合,即 O 位糖化,这种糖化需要在蛋白质中有一段特殊序列并处于折叠肽链的表面和对 GalNAc-转移酶作为受体的适宜环境。N 糖化与肽链合成同时进行,而 O 糖化在翻译后加工过程中可能与肽链折叠同时进行。小肽链短,所以糖化多肽并不多见,ACTH 在裂解前可能被糖化。

有些多肽前体由同一种基因在不同细胞表达,但相同前体的加工却有细胞特异性,在不同的组织中,同一种前体产生不同的活性多肽,如胰升血糖素前体在胰岛 α 细胞产生胰升血糖素,而在肠 L 细胞产生包含胰升血糖素序列在内的胃分泌抑制肽(oxynotomodulin)。说明不同细胞中加工胰升血糖素前体的酶系不同。

前体的加工场所主要在分泌颗粒,有的多肽在高尔基体开始加工,在分泌颗粒继续肽链翻译后加工,形成活性多肽。然后降解成无活性片段或具有其他生物活性的小片段。译后加工到降解是一个连续的过程,这个过程中最令人关注的是活性多肽的半生期,半生期长短不仅取决于多肽的结构,而且与该肽有无结合蛋白有关,已经发现不少多肽有结合蛋白,如 IGF、GH、CRF 等。与结合蛋白结合的多肽,其生物活性低但半生期延长。多肽降解或灭活实际上的重

要性可与多肽的合成相提并论,如果一个多肽在传递信息后应迅速灭活,但却继续存在,必将引起严重后果。

有时由于基因产生点突变使编码的氨基酸发生改变。如果氨基酸置换发生在重要位置,就可能造成严重后果。如胰岛素原 A 链或 B 链与连接肽的结合点上产生突变,使胰岛素原不能被酶切而使胰岛素产生减少。有遗传缺陷的 Brattleboro 大鼠不能编码加压素,是因为其外显子缺少一个碱基,在转录时缺少一个三联密码,导致合成与加压素序列不同的无活性肽链。

近年来发现一些多肽的降解产物在调节多肽合成分泌的自我反馈机制中起重要作用,如内源性促性腺激素释放激素 GnRH1~5 可抑制 GnRH 分泌,而 GnRH2~10 无此作用, C 肽可抑制胰岛素的分泌。而且属于不同的多肽的内源性降解产物也可能起到多肽分泌的调控作用,如 GnRH 的分泌涉及 N-甲基-D-门冬氨酸(NMDA)受体的激活,一些内源性多肽片段,如以谷氨酸残基为末端的 GRF1~37、IGF-1 1~3 和 C 肽作为 NMDA 受体拮抗剂而抑制 GnRH 的分泌,但它们的生理意义有待证实。mRNA 前体的不同拼接产生不同的多肽前体和肽链翻译后加工,使一个基因产生多种多肽,这是蛋白质或多肽合成的一个重要的进化。可以设想由于这种进化,使一定数量的基因产生多得多的活性多肽而能满足体内的生理需求。有的多肽前体中包含了几个序列相同的多肽和一些其他活性多肽,如脑啡肽前体含 4 个蛋-脑啡肽、2 个 C 端延长形式的蛋-脑啡肽和一个亮-脑啡肽, TRH 前体中有 6 个 TRH 直接前体 (TRH-GLY) 和 7 个插入肽,它们都可能具有生物活性,这种有重复序列的前体等于提高了基因表达的效率。

三、多肽的种间差异和同功能多肽的分子间差异

多数活性多肽分子本身或其前体的序列常有种间差异(表 1-4),由于这些肽链中氨基酸的差异多半不在分子中保守的关键性位置,因此多肽活性的种间差异并不如此明显。在临床应用中,猪与人胰岛素只有一个氨基酸不同,所以在用基因工程生产人胰岛素之前就被用于治疗。多肽序列的种间差异虽不明显影响生物活性,但长期应用不能排除产生抗体的可能性。所以在多肽生物活性研究中,不总是特别强调这种种间差异。但个别多肽的关键性位置的氨基酸种间差异可明显影响与异种动物的受体结合。如大鼠肾上腺髓质素(adrenomedullin)前体中的 ProAM 22~41 肽能明显增加大鼠心输出量,对其他动物无效。而人 ProAM 22~41 肽对试验过的动物也无作用。这个 20 肽中其实只有 3 个氨基酸不同。大分子多肽如垂体前叶激素肽链长,种间的序列差异的绝对量大,常造成对异种动物受体结合的明显影响。如低等动物的 GH 对人无效,就是由于这种明显差异所致。GH 具有很强的免疫原性,所以对相对分子质量较大具有三级结构的多肽种间差异是一个重要现象。

体内功能上基本相同而序列上也基本相同的多肽往往由不同的基因表达,如均为 37 肽的 CGRP α 、 β 二型有 3 个氨基酸的不同,但分别由 2 个不同的基因表达。垂体 GH 和胎盘 GH 的 191 个氨基酸中只有 13 个氨基酸不同,却各自由 GH-N 和 GH-V 基因表达,所以序列中某些位置的氨基酸不同反映了其基因本身结构的不同,与翻译后加工无关。但是当基因发生点突变后,可能导致编码的氨基酸不同。有些多肽前体序列有所不同,可能由于基因转录后前体 mRNA 经过不同的拼接产生不同 mRNA 所致。

表 1-4 多肽氨基酸序列的种间差异

名 称	氨基酸(人)	与其他动物的差别
促肾上腺皮质激素(ACTH)	39	大鼠 3, 猪 3, 牛 5, 羊 6
降钙素(galitonin)	32	鸡 2, Salmon 11
降钙素基因相关蛋白(CGRP)	37	鸡 4, 大鼠 3
心钠素(ANP)	28	大鼠 1
淀粉不溶素(amylin)	37	大鼠 6
脑钠素(BNP)	32	大鼠 16
促肾上腺皮质激素释放因子(CRF)	41	牛 7, 大鼠 0
内皮素(endothelin)	21	大鼠, 小鼠, 猪, 牛, 狗 0
促性腺激素释放因子(GRF)	44	牛 5
神经肽(NPY)	36	大鼠 0
肾上腺髓质素(adrenomedullin)	52	大鼠(50 肽)6
肾上腺髓质素原 N 端 20 肽(PAMP)	20	大鼠 3
胰岛素(Insulin)	51	猪 1, 牛 3
催产素(oxytocin)	9	哺乳动物 0
加压素(vasopressin)	9	哺乳动物 0 (猪 LYS VP)

四、多肽构效关系

有些多肽在序列上有相当高的同源性,虽然它们都是不同基因的产物。人们常将序列上有同源的不同多肽归为一个家族甚至超级家族。从构效关系考虑,序列上的同源性必然带来功能上的关联,如肾上腺髓质素的 C 端与 CGRP 有 9 个氨基酸相同,都有一个分子内二硫键,它们都有扩张血管、降低血压的作用。又如人 GH 和人 PRL 结构上有一定关联,人 GH 有催乳作用,PRL 也有促生长作用。但是多肽分子中少数关键位置的氨基酸不同却可呈明显不同的生物活性,下面是催产素和加压素序列的比较。

催产素: C-Y-I-Q-N-C-P-L-G-NH₂

加压素: C-Y-F-Q-N-C-P-R-G-NH₂

这 2 种垂体后叶激素都是 9 肽,分子中只有 3 位和 8 位氨基酸不同,虽然生物活性互有交叉,但差别非常明显。加压素主要作用是抗利尿、收缩血管、升高血压,而催产素主要引起子宫平滑肌收缩。

从加压素和催产素的例子可以看出多肽结构和功能的重要关系。ACTH 是 39 肽,ACTH_{1~24} 无种间差异,保持了 1~39 位原有生物活性所必需的最短肽链;而 ACTH_{1~17} 只有完整 ACTH 分子的活性的 5%。N 端去掉 1 位丝氨酸或丝氨酸被乙酰化,可使 ACTH 活性明显丧失,但 1 位丝氨酸被甘氨酸取代,不影响活性。(Gly¹-Lys¹⁷)-ACTH_{1~17} 却具有十分明显的促肾上腺皮质分泌的作用。胰岛素的 A、B 链共 51 个氨基酸,其中 29 个可被置换,不能置换的氨基酸中多数为带疏水侧链的非极性氨基酸,这些氨基酸和二硫键是维持胰岛素空间构型所必需的,若将 B 链 Lys 30-Thr³¹ 倒置成 Thr³⁰-Lys³¹,就能使胰岛素分子不易聚合。这些例子说明多肽的结构是决定多肽功能的基础,但是多肽分子结构中,一部分氨基酸对生物活性并非必需,它们可被切掉而不影响活性,而另一部分氨基酸是保持生物活性所必需的关键部分,常为几个甚或近全长的肽链。关键位置的氨基酸对活性特别重要,但是这部分氨基

酸的数目有可能被减少或被置换,在多肽链这2种情况都很难预测,只有通过结构和功能关系的研究才能有所了解。到目前为止,许多多肽的构效关系并不十分了解,因此,多肽构效关系的研究是多肽药物化学的重要研究课题。

关于多肽活性测定方法需视不同多肽而定,但常用超生理剂量来研究它们的活性,无论体内实验,常以浓度为 10^{-6} ~ 10^{-8} mol/L为宜,但实验室之间结果的比较有时有所不同,因为多肽纯度即使100%,也可能因种种原因使多肽失活而影响结果,所以用放射免疫测定只能定量,而不能反映多肽活性。结合生物活性而使多肽定量标准化十分困难,临床上应用国际单位表示活性多肽激素的含量,不仅要求应用国际标准的多肽激素标准,而且需经常核实生物活性是否保持不变。

结构研究的目的在于了解多肽分子功能的结构基础,以能有效地与受体结合,合成更强的生物活性的激动剂或合成仅能占据受体、本身对原多肽不产生效应的拮抗剂,由于多肽是多功能的生物活性物质,改造多肽分子有可能增强或减低某种功能,如TRH类似物DN-1417具有明显的中枢作用,而释放激素活性仅为TRH的1/40。

到目前为止,仅由10个氨基酸组成的GnRH类似物已超过2000种,但仅有1%~2%的类似物比较满意,说明类似物的合成尚无可靠的规律,多肽结构改造必需在生物活性监测下进行,生物活性测定是判断多肽结构改造成败的唯一标准。

研究多肽构效关系的常用方法大致可归纳为以下几个方面。

1. 确立保持完整多肽所具有的生物活性所需的链长

如GRF为44肽,研究构效时合成各种链长(N端完整)的GRF,然后与44肽比较生物活性。研究表明GRF只需N端29肽即可保持44肽同等水平的生物活性。一般说来,多肽N端常是多肽活性的结构基础,在序列上常比较保守,能满足受体结合的需要。但是少数多肽切去N端头几个氨基酸后,生物活性明显增加,如GLP-1(7~36)NH比GLP-1(1~36)NH₂活性明显增加。P物质去掉N端后,活性也增加。多肽分子C端的差异可能归结于种属特异性,如ACTH。有些多肽具有分子内二硫键形成的环形结构,环内氨基酸数也有可能减少或被置换,如SS的环内外氨基酸不仅减少且被置换而生物活性明显增强,SM201~995已用于临床。但有些具有二硫键的多肽分子当二硫键被打开时,使分子构型明显改变而失去生物活性。

2. 氨基酸置换

多肽分子中必需的结构确定后,可试用某个(些)氨基酸取代肽链中的某个(些)氨基酸残基,用D型氨基酸取代L型是常用的手段。氨基酸残基的替换有二个目的,一是与受体结合能力加强,二是使某个肽键稳定性增加,从而延长多肽半生期,如SS中Trp-Lys的肽键易被水解,将Trp换成D-Trp,则使SS半生期明显延长。octreotide是SS八肽类似物,如Thr8被Trp取代,可增加受体的亲和力,而Tyr3和Val6对GH分泌的抑制更加特异。

3. 引入氨基酸衍生物或修饰氨基酸侧链

这是常用的方法。如GnRF5位Gly换成D-Nal(2)即3-(2-萘基)-D-丙氨酸后,类似物活性比天然GnRF高200倍以上,临床上应用的抗利尿激素dDAVP[1位为半胱氨酸去氨基(3-mercaptopropionyl)和8位D型精氨酸],它们抗利尿和加压作用的比值比加压素大2000~3000倍,半生期也明显延长。