

综述与专论

设计形成特定构象单元的多肽

徐梅 鲁子贤

(中国科学院上海生物化学研究所上海, 200031)

提 要

肽分子设计的命题,基本思想是以特定的构象单元与功能的关系为指导,设计具有特定功能的多肽,它不仅可用于进一步揭示天然蛋白质的结构原则及卷曲机制,而且可以指导药物合成及工业产品的设计。本文旨在介绍一些设计形成特定构象单元的多肽的进展,并从中归纳出一些设计原则。

关键词 多肽设计,设计原则

肽的研究是近十年来分子生物学发展最快的领域之一,残基置换是长期来研究肽的结构与功能的关系及寻找高效药物的重要手段,其置换原则虽已归纳出一些规则,但其实质是经验性的,而且变化的范围限于单个或局部的氨基酸。随着对肽和蛋白质的构象单元卷曲机制,结构(含立体结构)与功能的关系的深入研究,对已知结构的蛋白质的数据库的全面分析,及构象预测理论不断进步,逐步提出了肽分子设计的命题,至目前止,基本的设计思想是以特定的构象单元与功能的关系为指导,设计具有特定功能的多肽,这正是分子设计与经验置换的本质不同。它利用已知的蛋白质结构原理,从设计结构的角度出发,获得新的多肽,使其具有与天然多肽相似或更特定的结构,并使其具有人类意志所要求的功能。这样一种途径,不仅可以用于进一步揭示天然蛋白质的结构原则及卷曲机制,完善多肽的设计原则,而且可以指导药物合成及工业产品的设计,因此它既是一种基础研究的手段,也是应用研究的重要方法。

由于蛋白质工程和基因工程的迅速发展,使得多肽设计的内容十分丰富,本文旨在介绍

一些设计形成特定构象单元的多肽的进展,并从中归纳出的一些设计原则。

一、设计原则

1. 三级结构的设计

正如房屋的设计需首先考虑总体的结构,多肽设计的第一步是主链框架即三级结构的设计。

肽链的长度与形成稳定的构象的关系经过长期的讨论,目前一般接受的观点为至少由50—60个残基组成的肽链才可能具有稳定的三级结构。而目前的多肽合成技术限制了合成肽链的长度,一般限于100个氨基酸残基,因此三级结构的设计目前大多是一些小而稳定的三级结构单元。例如比较常见的三种结构单元,上下叠垒(up-and-down β -barrel),希腊钥匙叠垒(Greek key β -barrel),上下四 α -螺旋簇(up-and-down four α -helix cluster)^[1]它们需60—80个残基即可稳定,其结构简单,具有普遍性,研究这些结构,从中可以得到许多通用的结构原则。

当然,如果选择的结构单元不含三级结构,则可以直接从二级结构入手。

2. 序列的设计

三级结构是二级结构装配而成, 三级结构选定后, 就必须进行二级结构和肽链转角的设计, 使选择的序列趋于形成所设计的三级结构, 在这里考虑各种相互作用力对多肽结构的稳定作用是重要的。

(1) 转角的设计: 对各种转角的研究表明, 回折结构不仅决定了肽链的走向, 而且位于多肽分子表面, 直接与环境相作用, 对稳定多肽的结构起重要作用, 为了使上下螺旋或折叠紧密相联, 转折点的位置必须满足肽链走向和氢键的分布。特定位置转角(1—4)的氨基酸残基一般为 2-Pro, 1or 3-Asn, 根据统计, 蛋白质中 2/3 的 Pro-Asn 为转角的中心^[2]。此外, Gly, Ser 也是转角好的形成者。

(2) 二级结构的设计: 利用 Chou-Fasman^[3], Ptitsyn-Finkelstein^[4] 等方法可以预测假定序列的 α -螺旋和 β -折叠的形成势, 对于设计的 β -折叠, 序列的预测应为 $P\beta$ 高 $P\alpha$ 低, 对于设计的 α -螺旋, 则应是 $P\alpha$ 高 $P\beta$ 低。

(3) 疏水相互作用: 疏水相互作用对稳定蛋白质的结构起重要作用, 利用螺旋轮作图可以预测蛋白质中的两亲 α -螺旋片段^[5], 而亲疏水氨基酸相间组成的序列则可以形成两亲的 β -折叠。

(4) 空间的密堆积: 一些侧链较小的氨基酸, 虽然不是好的 α -螺旋或 β -折叠的形成者, 如 Gly, Ala, 但对于蛋白质内部的密堆积有好处时, 可考虑选择。

在稳定的三级结构中, 相邻的氨基酸残基的侧链非常靠近, 应考虑残基侧链在电荷、疏水、体积上的匹配, 如支链的残基不应相邻^[6]。

(5) 从研究结构角度出发, 选择的序列应与天然结构的序列具有最小的同源性, 如果同时研究其功能时, 其结合位点必须保留。

(6) 由于酰胺键的偶极性, 使得 α -螺旋 β -折叠的链呈偶极性, 从 C 指向 N, 因此在设计时, 应考虑到反平行的二级结构有利于偶极相互作用^[7]。

(7) 带异性电荷的残基空间距离 $< 4 \text{ \AA}$

时, 形成离子对^[8], 这种静电相互作用在设计酶与底物或激素与受体结合等分子与配体相互作用时显得特别重要。

(8) 立体匹配: 利用计算机建立蛋白质分子模型, 研究二级结构之间的相互作用, 适当的进行氨基酸置换, 使分子内部的堆积最适, 能量达到最优^[9]。

(9) 二硫键的引入会增加蛋白质的稳定性, 但同时可能破坏其它稳定蛋白质的作用力, 因此需借助已了解的三维结构或分子模型, 选择合适的二硫键的位点。

由于蛋白质的多样性, 具体设计时, 还应考虑到各蛋白质的特点。

二、设计实例

目前, 关于形成特定构象单元的多肽的设计工作很多, 大致分为两类, 基于结构的从头设计和基于功能的模拟设计, 本文将介绍这两类工作的一些最新进展。

1. 从头设计

(1) 4- α -螺旋簇的设计^[10]

细胞色素 b562, 细胞色素 c', 蚯蚓血红蛋白, 烟草镶嵌病毒蛋白, T4 噬菌体溶菌酶等共同含有相连的 4- α -螺旋簇, 如图 1 所示, 它们之间功能不相同, 序列也不相似。

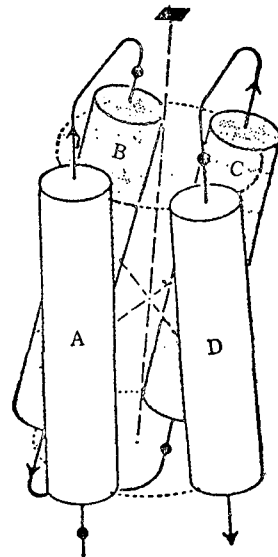


图 1 4- α -螺旋簇模型图

在螺旋簇中,肽链呈右手螺旋,三个连续的氢键网却形成左手螺旋,共同组成如电缆般缠绕的螺旋,其结构非常稳定,CO和NH的排列与肽链平行,使得整个螺旋因偶极矩而呈极性,而相邻的反平行的螺旋则有利于螺旋偶极相互作用。

不同的蛋白质中的螺旋的内部连结区的大小和构象差异很大,这表明螺旋是4- α -螺旋簇蛋白质卷曲的决定因素,那么如何设计最小的螺旋连结,细胞色素c的晶体结构表明,仅一个103-Gly就连接了相邻螺旋的末端102-Ala和104-Pro,固定了螺旋间的相互转动,而且二螺旋间的一组Ala相互作用使二螺旋紧密堆积。

在考虑了上述因素后,就可以选择具有强 α -螺旋形成势的氨基酸序列,同时使得螺旋间有稳定的疏水相互作用,还可以选择带电荷的残基,与螺旋的微偶极形成电荷-偶极或偶极-偶极相互作用。

(2) 上下叠垒的设计^[4]

Betaballin为一完全人工设计的上下 β -叠垒,如图2所示,它由二个相同的 β -折叠(31a,a)组成,每个 β -折叠由四个 β -链组成, β -链最少由6个残基组成。为了满足肽链的走向和氢键的正确形成,b-链在8位转向,转角的残基为Pro-Asn,考虑到疏水相互作用,肽链由亲疏水氨基酸残基相间而成,再根据二级结构的

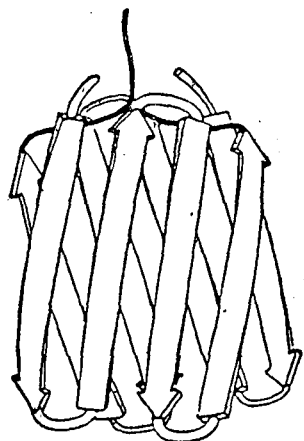


图2 上下叠垒模型图

预测,选择了一些强 β -形成者,如Val,Leu,Ile,Phe,每个链含四个强 β -形成者,Ala虽然不是强 β -形成者,但因其很小的侧链可以使蛋白质内部紧密堆积,因此,也可以选用,由于含支链的残基是强的 β -形成者,在序列中经常出现,为了蛋白质结构的紧密堆积,应使它们不要相邻。

点面图(Dot-surface computer graphics)表明:二个相对的 β -折叠的交角应为30—40°相邻的 β -折叠的接触多于相交的 β -折叠,CPK model的计算则建议12-Phe换成Tyr,以Tyr的OH代替Phe的环与溶剂接触,1-Thr换成Ser,因为1-Thr和2-Pro在发夹回折处阻碍了密堆积。

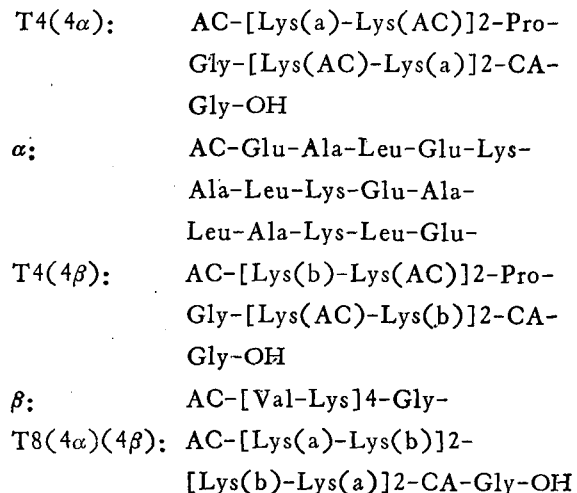
再结合其它计算模型,得到最后的序列为:
STVTARQPNVTYSISNPNTATVKLPNYL-
SIG

CD的结果表明,该模型肽含40%的 β -结构。

(3) 模板装配蛋白的设计^[10]

模板装配蛋白是将事先设计并合成好的带有结合位点的各类小肽采用化学缩合的方法,接入事先设计并合成好的带有相应结合位点的U-形回折(即模板)上所得到的全新的蛋白质。

首先合成了模板T和形成 α , β 结构的小肽,其序列如下



模板T中的Lys是为了形成 β -结构,Pro-Gly形成回折,整个模板形成发夹结构, β 中疏

水的 Val 和亲水的 Lys 交替, 可以形成两亲的 β -折叠, α 的 Glu Lys Leu Ala 均为高 α -螺旋形成者, 其排列顺序不仅可以形成两亲的螺旋, 而且 i 位和 $i + 4$ 位的残基间还存在静电相互作用, 同时带电荷的残基也利于螺旋内部的极性相互作用。

利用化学合成的方法, 将小肽接在模板的特定位置上, 就可以得到 $T4(4\alpha)-4\alpha$, $T4(4\beta)-4\beta$, $T8(4\alpha)(4\beta)-4\alpha 4\beta$, 如图 3 所示。

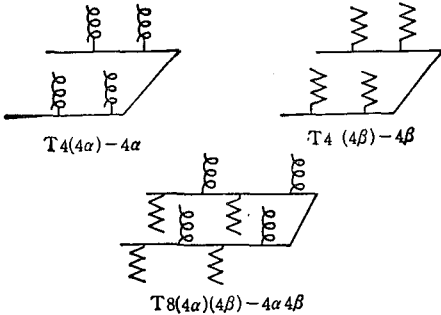


图 3 模板装配蛋白模型图

CD 结果表明: $T4(4\alpha)-4\alpha$, $T8(4\alpha)(4\beta)-4\alpha 4\beta$ 都形成 α -螺旋, $T4(4\beta)-4\beta$ 在 $pH = 1, 2$ 时形成 β -折叠。

2. 模拟设计

(1) 两亲肽的设计

两亲的二级结构在许多肽和膜结合蛋白质的生理活性中起重要作用。已有许多的计算机程序用于寻找可能的两亲片段, 最简便的是作图法, 上文已作了介绍。

两亲肽的设计基于如下假设: 如果肽或蛋白质的生理活性是依赖于某一区域的特定的二级结构, 那么就建立模型, 此模型与天然区域具有相近的二级结构形成势而其氨基酸序列却是最小同源, 然后考察模型的性质。

① 两亲螺旋的设计

根据分子模型推测 ApoA-1 含有 22 个残基的两亲螺旋, 其终止点为 Gly 或 Pro, 据此建立了 Model ApoA-1^[11] 由 Leu (疏水的 α 形成者), Glu (亲水带负电), Lys (亲水带正电) 交替重复组成, 具有两亲螺旋的形成势, 该螺旋含 1/3 的疏水面, 此序列与 ApoA-1 中的

两亲区最小同源, 22 个残基的 Model ApoA-1 与 243 个残基的 ApoA-1 相比, 同样具有结合磷脂, 在水/三氟乙醇中形成 α -螺旋, 及聚集行为, 此外, 还有 ApoA-1 的激活磷脂酶等生理功能。

② 两亲 β -链的设计

ApoB 在与脂蛋白质的表面结合时, 形成部分两亲 β -链, 根据两亲折叠的形成原则, 模型 ApoB^[11] 由 Val-Glu-Val-Orn 重复组成, 亲疏水残基交替, 亲水面上酸碱性残基交替, 模型 ApoB 达到了 ApoB 所有的重要的生理功能。

此外, 如蜂毒素, 调钙蛋白, 降钙素^[12] 的设计也获得了成功。

(2) 膜通道的设计

DeGrado 等^[13] 设计合成并研究的序列 (LSLLLSL) 3 和 (LSSLLSL) 3 以螺旋的形式与膜结合, 而且在有电压的情况下, 表现出天然离子的导体性质, 计算机所建立的模型表明, (LSLLLSL) 3 形成垂直于膜的四体螺旋, (LSSLLSL) 3 则形成以 Ser 为中心的六螺旋。

(3) 纤维结构的设计

许多纤维状蛋白质是由重复序列的长链绞合成的单股或多股的螺旋, 如胶原蛋白是由重复的 Gly-Pro-X 的序列形成的伸展螺旋, Darmstadt 等合成了三股链的胶原蛋白的类似物^[14]。

Tropomyosin 是由重复的七股链组成的螺旋卷曲而成, adges 等^[15] 成功地用模型肽 (LEALEGK) 5 模拟了这一结构。

弹性蛋白的弹性依赖于肽链伸展时的熵减, 根据弹性蛋白中许多重复的五肽而设计的弹性螺旋模型^[16] 表现了弹性蛋白的行为。

(4) 甜蛋白的设计^[17]

Monellin, Thanmatin 均为甜蛋白, 前者含 A(45aa), B(50aa) 二链, 不含二硫键, 对热不稳定, 后者为单链含 8 对 S-S 键, 对热稳定。

为了使 Monellin 同样具有热稳定性, 受

(下转第 356 页)

的不稳定性和激活的蛋白 C、凝血酶引起的 PAI 失活, 可能使 PAI 逐渐从血凝块中消失, 随后 PA 扩散进入血块, 导致血块水解。血浆中 PAI 活性可能在 PA 活性的系统控制中起作用, 防止在“休息”状态下 PA 起作用。

此外, Hill 等人发现, 丝氨酸蛋白酶抑制剂是防御入侵寄生物的有效物质。PAI 除在凝血、纤溶等过程中起关键性调控作用外, 是否还象其它丝氨酸蛋白酶抑制剂那样在机体防御系统中起重要作用, 这一点还有待于进一步研究。

参 考 文 献

- 1 Sprengers E D, Kluft C. *Blood*, 1987; 69(2): 381
- 2 Hart D A, Rehemtulla A. *Comp Biochem Physiol*, 1988; 90B(4): 691
- 3 Wiman B, Chmielewska J, Rånby M. *J Biol Chem*, 1984; 259(6): 3644
- 4 Andreasen P A, Riccio A, Welinder K G *et al. FEBS Lett*, 1986; 209(2): 213
- 5 Levin E G. *Blood*, 1986; 67(5): 1309
- 6 Levin E G, Santell L. *Blood*, 1987; 70(4): 1090.

- 7 Hekman C M, Loskutoff D J. *J Biol Chem*, 1985; 260(21): 11581
- 8 Lambers J W J, Cammenga M, König B W *et al. J Biol Chem*, 1987; 262(36): 17492
- 9 Nielsen L S, Andreasen P A, Grøndahl-Hansen J. *FEBS Lett*, 1986; 196(2): 269
- 10 Loskutoff D J, Linders M, Keijer J *et al. Biochem*, 1987; 26(13): 3763
- 11 Zehnb R, Gelehrter T D. *Gene*, 1988; 73(2): 459
- 12 Wun T C, Reich E. *J Biol Chem*, 1987; 262(8): 3646
- 13 Hibino T, Izaki S, Ohkuma M *et al. FEBS Lett*, 1988; 231(1): 202
- 14 Chapman H A Jr, Stone O L. *Biochem J*, 1985; 230(1): 109
- 15 Wohlwend A, Belin D, Vassalli J D. *J Immunol*, 1987; 139(4): 1278
- 16 Kiso U, Kaudewitz H, Henschen A *et al. FEBS Lett*, 1988; 230(1, 2): 51
- 17 Ye R D, Wun T C, Sadler J E. *J Biol Chem*, 1987; 262(8): 3718
- 18 Stump D C, Thienpont M, Collen D. *J Biol Chem*, 1986; 261(3): 1267
- 19 Scott R W, Bergman B L, Bajpai A *et al. J Biol Chem*, 1985; 260(11): 7029
- 20 Chmielewska J, Rånby M, Wiman B. *Thromb Res*, 1983; 31(3): 427

[本文于1990年7月9日收到, 10月9日修回]

(上接第 328 页)

Thanmatin 单链的启发, 设计了 link (序列为 YENERIK), 将 B46-Ile 与 A6-Gly 联接起来, 得到单链的甜蛋白 Model-Monellin, 对热及酸碱稳定。

三、多肽设计的前景

蛋白质化学已进入蛋白质工程的时代, 多肽的设计成为其重要的前提, 目前的工作已初步取得令人兴奋的结果, 各实验室的工作不仅在上述工作的基础上进一步发展, 而且不断提出各种新颖的设计思想^[18], 随着多肽合成和基因合成技术的进步, 将使合成大而完整的蛋白质成为可能, 分离方法的突破, 将使多肽的量满足结晶和核磁的研究用量。从各个角度进行多肽的设计不仅有助于丰富已有的蛋白质的结构知识, 而且激发我们以崭新的观点去考察蛋白质的结构原理和卷曲机制。

参 考 文 献

- 1 Jane S R *et al. Protein Engineering*. Alan R Liss, Inc,

- 1987: 149—163
- 2 Zimmerman S S *et al. Proc Natl Acad Sci USA*, 1977; 74: 4126
- 3 Chou P Y *et al. Biochemistry*, 1974; 13: 222
- 4 Ptityn O B *et al. Biopolymer*, 1983; 22: 15
- 5 Schiffer M *et al. Biophys J*, 1967; 7: 121
- 6 Crippen G M, Kuntz I D. *Int J Pept Protein Res*, 1978; 12: 47
- 7 Douglas H O *et al. Protein Engineering*. Alan R Liss Inc, 1987; 165—173
- 8 Boldwin R L *et al. Protein Engineering*, Alan R Liss Inc, 1987: 127—148
- 9 Connolly M L. *Science*, 1983; 221: 709
- 10 Manfred M *et al. Proteins: Structure, Function, and Genetics*, 1989; 5: 13
- 11 Kaiser E T. *Protein Engineering*. Alan R Liss Inc, 1987: 193—199
- 12 Susan E V *et al. Protein Engineering*. Alan R Liss Inc, 1987: 201—211
- 13 DeGrado G F *et al. Science*, 1988; 240: 1177
- 14 Germann H P, Heidemann E. *Biopolymer*, 1988; 27: 157
- 15 Talbot J A, Hodges R S. *Acc Chem Res*, 1982; 15: 224
- 16 Urry D W. *J Protein Chem*, 1988; 7: 1
- 17 Kim S H *et al. Proteins: Structure, Function, and Genetics*, 1989; 2(8): 571
- 18 Richardson J R *et al. TIBS*, 1989; 14(7): 304

[本文于1990年7月12日收到, 10月5日修回]