

蛋白质空间结构中的“超二级结构”“模体”和“结构域”概念浅析

张海涛*

(广东医科大学生物化学与分子生物学教研室, 湛江 524023)

摘要: 生物化学教材中关于蛋白质结构中的“超二级结构”“模体”和“结构域”的概念存在着一些模糊区域。不同的教材对这些概念解释不一, 甚至是相互矛盾的, 导致学生理解这些概念相对困难, 是教学的难点。作者根据蛋白质空间结构定义和教学经验, 对这些概念进行对比解析, 提出理清这些模糊概念的基本方法。该文旨在促进这几个模糊不一的概念能在教学和研究中统一描述, 便于教学和学术交流。

关键词: 二级结构; 超二级结构; 模体; 结构域; 三级结构

A brief analysis of the concepts of "supersecondary structure", "motif" and "domain" in protein space structure

ZHANG Haitao*

(Department of Biochemistry and Molecular Biology, Guangdong Medical University, Zhanjiang 524023, China)

Abstract: There are some fuzzy areas in the description of "super-secondary structure", "motif" and "domain" of protein structure in the current biochemistry textbooks. It is difficult for students to understand those concepts, because the description of these concepts is inconsistent, even contradictory in the different textbooks, which is also the teaching difficulty. According to the definition of protein structure and teaching experience, the author makes a comparative analysis of these concepts and puts forward a basic method to clarify these fuzzy concepts. The purpose of this paper is to promote the unified description of these concepts in teaching and research, and facilitate teaching and academic exchanges.

Key Words: secondary structure; supersecondary structure; motif; domain; tertiary structure

在讲解蛋白质的结构功能时, 为了更好地解析这类复杂的生物大分子结构, 蛋白质结构一般被人为地分为一级结构、二级结构、三级结构和四级结构进行逐层解说。教材对于蛋白质这几层结构的定义基本是统一的, 但是随着超二级结构、模体和结构域的概念被引入教材后, 不同教材对二级结构、超二级结构、模体、结构域和三

级结构的定义和关系的描述和解释出现了差异, 使得原本简单明了结构解释变得模糊。这些模糊而不同的解释对于刚接触生物化学课程的学生来说是抽象且不好理解的。笔者根据教学经验和学习文献, 比较不同的教材对于这些概念解释和分类的合理性, 提出一些解决这一模糊区域的方法和分类, 以供从事生物化学教学的同行分析和商

收稿日期: 2022-06-09

基金项目: 广东省本科高校教学质量与教学改革工程建设项目(粤教高函[2021]29号)

*通信作者: E-mail: taohaizhang33@163.com

榘, 从而促进这几个模糊不一的概念能在教学和研究中统一。

1 二级结构定义及其基本类型

我们首先要明确蛋白质的二级结构和三级结构的基本概念。二级结构和三级结构均属于蛋白质空间结构, 二者的区别在于是否涉及描述组成蛋白质分子的氨基酸残基侧链基团的空间构象^[1-6]。其中不同的教材对于蛋白质三级结构的定义和解释是统一的, 但是教材关于蛋白质二级结构的定义是一个发展变化过程^[1-6]。

为了更好理解组成蛋白质的肽单元空间构象, 研究者根据肽单元空间构象的差异, 提出蛋白质二级结构的概念, 因此二级结构关注的是肽单元的空间结构形式, 而不考虑组成这一区域氨基酸残基侧链的构象。早期的生物化学教材关于二级结构概念是指的蛋白质肽链中肽单元局部的空间走向, 限定了二级结构的区域是局部^[1,5], 在教材中又列举了 α -螺旋和 β -折叠作为例子详细介绍, 最终导致了二级结构的概念是局部肽单元结构的局限认识, 这也导致早期教材将二级结构和二级结构类型的概念混淆^[1,5]。由于蛋白质三级结构描述的是蛋白质所有组成分子的空间构象, 为了与蛋白质三级结构的概念对应, 笔者认为蛋白质的二级结构描述为“蛋白质多肽链上的所有肽

单元在空间排列走向”^[3]比“局部的肽单元空间排列走向”^[1]更合理, 也更符合逻辑性。

肽单元上 C_{α} 原子所连的2个单键旋转角度差异使得2个相邻的肽单元具有相对空间位置, 造成组成肽链的所有肽单元在空间走向上会形成不同的特定规律的结构类型。这些特定规律的结构类型依据肽单元的空间排列走向的形态特征被命名为 α -螺旋、 β -折叠片(或称 β -折叠)、 β -转角(环)等。除了这些有特定规律的结构类型外, 蛋白质分子中还存在有序非重复性且无确定规律性形状的肽单元构象, 称为有序无规律结构(ordered nonregular structures)或有序非重复结构(ordered nonrepetitive structures), 旧称为无规卷曲, 即泛指那些不能被归入明确结构类型的二级结构形状^[2,4,6]。这部分肽链的肽单元平面间的二面角的角度不一, 无统一的规律性, 但对于某种类的蛋白质分子而言, “有序无规律结构”具有明确而稳定的结构^[2,4,6]。一个蛋白质分子可包含1种或多种类型的规律二级结构类型和有序无规律结构, 它们组成了一个蛋白质分子的完整二级结构, 即所有肽单元在空间走向(图1)。

2 超二级结构是有规律的二级结构类型

随着对蛋白质空间结构的深入研究解析, 研究者发现蛋白质多肽链顺序上相互邻近的一些二

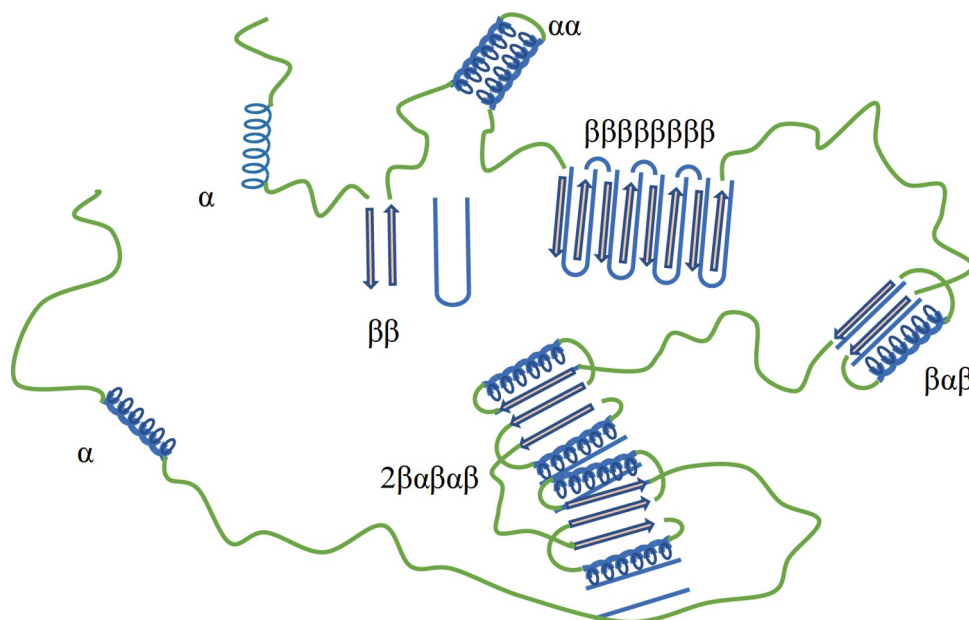


图1 蛋白质二级结构和二级结构类型的示意图

级结构类型会形成特定规律的聚集体,如 α -螺旋组合(α -螺旋- α -螺旋结构, $\alpha\alpha$), β -折叠组合(β -折叠- β -转角- β -折叠, $\beta\beta\beta$)和 α -螺旋 β -折叠组合($\beta\alpha\beta$)。这些特定二级结构类型的聚集体被称为“超二级结构”^[1,2],而实际上这些聚集体依然是蛋白质分子中有规律的局部肽单元空间走向,本质上仍属于一种有规律的二级结构类型,而不是所谓的介于二级结构和三级结构的一种特殊的空间结构。

我们通过对 β -折叠(β -pleated sheet或 β -sheet)的释义就很好理解为何“超二级结构”依然是二级结构类型。根据教材对 β -折叠的解释,我们知道稳定 β -折叠结构的氢键有2条平行的 β -折叠肽段形成的,所以至少要有2股平行的 β -折叠肽段才能使其稳定存在,而且肽段回折构成平行走向时,必然存在一个转角或环,理论上单股 β -折叠肽段是不能存在天然蛋白质分子中的。可见,原归于二级结构类型的 β -折叠天然就是一个 β 折叠组合(β -折叠- β -转角- β -折叠, $\beta\beta\beta$),完全符合现在对“超二级结构”的定义范畴。

完整的蛋白质二级结构不仅包括我们可以分辨出的有规律的肽单元结构,如螺旋、折叠、转角(环)等单位以及由这些单位相互组合的“超二级结构”,还包括大量有序无规律结构。如果以局部肽单元在空间的走向定义为二级结构,就会将 α -螺旋、 β -折叠等二级结构类型看成二级结构的概念,则必然会出现由多个二级结构集合组成的结构单位——“超二级结构”^[1,5]。如果二级结构被定义为一种蛋白质分子的所有肽单元在空间走向时,那就不会出现这个“超二级结构”的概念。所谓的“超二级结构”实际上也是一种蛋白质的局部肽单元在空间的走向,其本质仍然是一种有特定规律的二级结构类型。原来的“超二级结构”本意应该是比单一形态的规律性二级结构类型更为复杂的组合型规律性的超“二级结构类型”。因此,把“超二级结构”解释为是介于二级结构和三级结构间的一种特殊的空间结构的说法是令人费解的,并且容易打乱蛋白质分子结构的分级定义。

我们可以用组成一个蛋白质分子的所有肽单元在空间的走向(蛋白质二级结构)的示意图区别二级结构和二级结构类型(含超“二级结构类型”)

(图1)。图中细线条代表肽单元在空间没有规律的二级结构类型区域,即有序无规律结构区域,弹簧线圈代表 α -螺旋,宽箭头代表 β -折叠,肽链180°转向处是 β -转角(环)。这些不同的单一规律性二级结构类型可以单独存在,也可以与其他二级结构类型组合一个超“二级结构类型”,如 β -折叠- β -转角- β -折叠形成的 $\beta\beta\beta$ 结构类型, α -螺旋- α -螺旋形成的 $\alpha\alpha$ 结构类型,以及希腊图案样折叠形式 β -折叠- β -转角组合体($\beta\beta\beta\beta\beta\beta$)或罗斯曼(Rossmann)折叠模式 $\beta\alpha\beta\alpha\beta$ 。从图1中也可以看出,这些超“二级结构类型”实际上也只不过是结构更为复杂的有规律的二级结构类型,因此把它们看作介于二级结构和三级结构之间的所谓“超二级结构”就有些牵强了,实际上就不应该有一个“超二级结构”的概念^[6],最多有个“超二级结构类型”的概念。

3 模体是局部有规律的蛋白质三级结构类型

在对蛋白质结构的研究和认识过程中,研究者发现,在不同的蛋白质分子中具有相似的特殊结构类型区域可以执行类似的生物功能,于是将这些特殊结构类型命名为模体或基序(motif)。关于模体的概念在不同的教材也如“超二级结构”这个词一样,从开始就没有一个准确的定义,不同的教材解释也不一致,有的教材认为模体等同于“超二级结构”^[1,2],有的教材又认为模体与结构域等同。

模体这个名词在生物信息学研究中是指蛋白质分子中一段有特定规律折叠组成,并执行相应生物功能的保守多肽,其强调的是折叠和功能。由于三级结构才具备强调蛋白分子功能的基础,因此模体结构理应包括氨基酸残基侧链(包括辅助元素)的空间构象。笔者认为把模体看作是一个蛋白质分子的局部规律性的三级结构类型更为合理。如果把不同结构类型的模体看成蛋白质三级结构中有规律的三级结构类型(类似于二级结构类型,如 α -螺旋、 β -折叠片、罗斯曼折叠等),则对于蛋白质结构分层解释,会显得更清晰,更好理解和把握,因此笔者建议是否可以将模体定义为蛋白质分子中局部有规律的二级结构类型与组成这些结构类型的氨基酸残基侧链基团相互作用折

叠形成的结构功能单位,即蛋白质分子局部有规律的三级结构类型。

确定定义以后,我们就可以将模体看成是蛋白质局部二级结构和氨基酸残基侧链基团组成的空间聚合体,是局部规律性的三级结构类型(图2)。例如,锌指结构(zinc finger)是含有一个 $\beta\alpha\beta$ 组合形式的二级结构类型与这区域的氨基酸残基侧链(锌离子可以看做侧链的一部分)相互作用折叠组成的功能单位(图2A)。模体结构组成也可以是由不同亚基的肽链分子间组合而成,如亮氨酸拉链结构(图2B),即由两条肽链相邻而形成,这也意味这种模体类型是两个亚基局部区域组成。模体也可以由不是已知二级结构类型聚集形成,而仅由几个氨基酸残基组成,例如纤连蛋白识别结合受体的肽段,只是RGD三肽。可见,“超二级结构”可以形成模体结构,但是模体结构不一定存在“超二级结构”。

据此,我们可以将蛋白质二级结构类型和三级结构的类型逐一对应理解。不同的二级结构类型或聚集体加上组成这一区域的氨基酸残基侧链空间构象形成了有规律的三级结构类型——模体。有序无规律结构加上组成这一区域的氨基酸残基侧链空间构象形成蛋白质分子中无确定规律性的三级结构类型,可称蛋白质分子无定型结构区域。

4 模体与结构域的关系

蛋白质局部空间结构的描述中除了模体这一结构功能单位外,还有一个空间结构功能单位概念——结构域(domain)。结构域与模体之间的关系

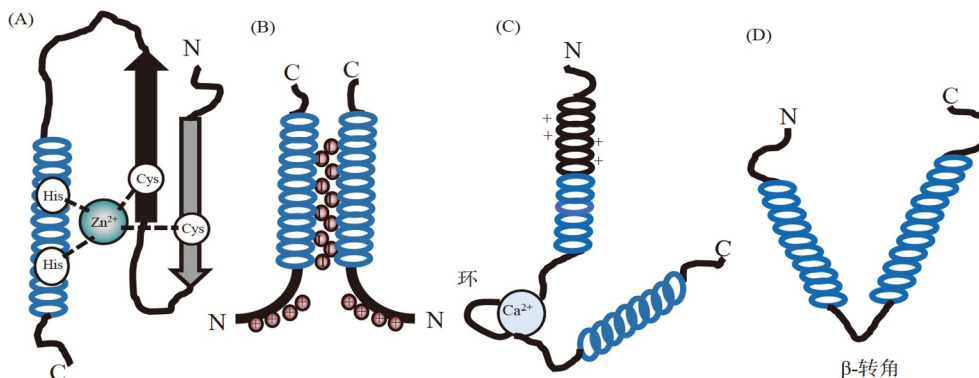
又如何呢?

现行的教材普遍将结构域定义为:蛋白质三级结构含有的多个分界明显、独立折叠、功能差异的球形结构^[1-6]。我们可以理解为蛋白质多肽链某些区域的肽单元形成有规则的二级结构类型,相邻的二级结构类型集装形成多个二级结构类型复合体,加上组成这一区域的各氨基酸残基侧链与肽单元相互影响,折叠成近似于球状结构的功能单位。这提示模体和结构域可以等同也可以不同。如果模体结构在蛋白质分子独立折叠为球形结构功能单位,则可视之为结构域,这时一个模体单位等同于一个结构域单位。如果几个模体单位组成大的球形结构域,则可视模体为结构域的一个组分单位,此时结构域由多个模体组成。可见,结构域是由一个或多个模体组成的蛋白质局部三级结构区域。将模体与结构域的关系类比于二级结构类型与“超二级结构类型”的关系可更好地让学生理解这两者的联系。我们也可以用这一规律认识到结构域和蛋白质三级结构的区别,即一个或多个结构域加上所有的无定形结构区域就形成了完整的蛋白质三级结构。

5 总结

为了更好推进对蛋白质空间结构的理解和交流,根据二级结构和三级结构定义上的差异,笔者尝试将涉及到其中的空间结构概念做了一些粗略的梳理,总结如下。

蛋白质二级结构是指组成一个蛋白质分子的所有肽单元空间走向,不含氨基酸残基侧链空间构象,而局部肽单元空间结构均属于二级结构类



A: 锌指结构; B: 亮氨酸拉链结构; C: 碱性 α -螺旋- β -环- α -螺旋结构; D: α -螺旋- β -转角- α -螺旋结构

图2 模体类型

二级结构范畴：
肽单元空间走向

三级结构范畴：
肽单元与氨基酸残
基侧链的空间结构

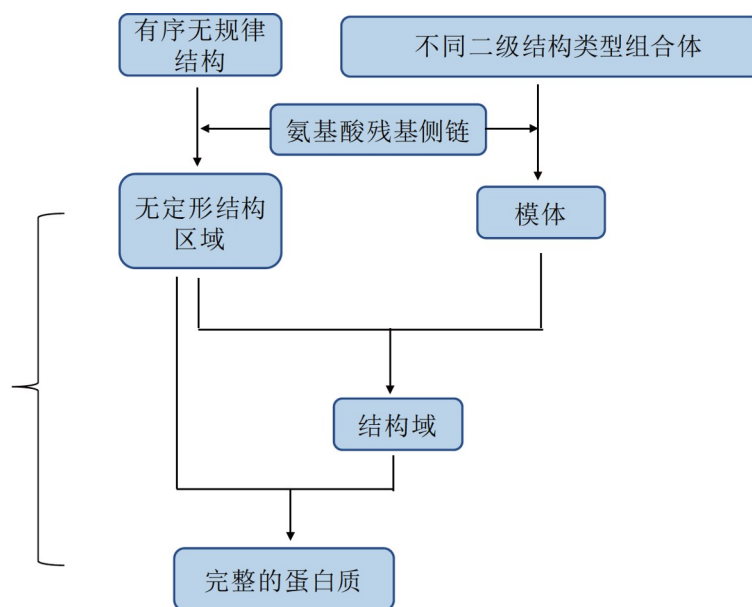


图3 二级结构单位和三级结构单位间的关系

型(包括“超二级结构”)。

一个蛋白质完整的空间结构是其三级结构，包括组成这个蛋白质所有肽单元和氨基酸残基侧链的空间结构。模体或结构域则是由不同的二级结构类型加上组成这一区域的氨基酸残基侧链的空间结构折叠组成的功能单位，是有规律的三级结构类型。蛋白质分子中可以由一个或多个模体结构组成一个球形的结构域，而一个或多个结构域加上所有的无定形结构区域则构成完整的蛋白质三级结构。二级结构单位和三级结构单位间的具体关系如图3。

进行这样的梳理不仅有助于避免教学中出现的模糊区域导致学生理解这些概念的困难，而且也有助于在蛋白质结构研究中促成准确统一规范

使用这些名词，为更好地进行教学以及学术交流等提供一定帮助。

参考文献

- [1] 周春燕, 药立波. 生物化学与分子生物学(第9版). 北京: 人民卫生出版社, 2018
- [2] 刘新光, 罗德生. 生物化学与分子生物学(第3版). 北京: 科学出版社, 2021
- [3] 赵炜明, 宋高臣. 生物化学与分子生物学(第1版). 北京: 人民卫生出版社, 2020
- [4] 陈娟, 李凌. 医学生物化学与分子生物学(第4版). 北京: 科学出版社, 2022
- [5] 王镜岩, 朱圣庚, 徐长法. 生物化学(第3版). 北京: 高等教育出版社, 2002
- [6] David L, Nelson, Michael M, Cox. Lehninger principles of biochemistry (8th ed). New York: W. H. Freeman, 2021