

学校编码: 10384

分类号: _____ 密级: 机密

学 号: B200026009

UDC: _____

学 位 论 文

戊型肝炎病毒衣壳蛋白体外组装相关结构域的研究

李 少 伟

指导教师: 陈睦传 教授夏宁邵 研究员厦门大学生命科学学院生物学系申请学位级别: 博 士 专业名称: 动 物 学论文提交日期: 2003年10月27日 论文答辩日期: 2003年11月11日学位授予单位和日期: 厦 门 大 学答辩委员会主席: 朱关福 研究员 (博导)评阅人: 朱关福 研究员 (博导)詹美云 研究员 (博导)戚中田 教授 (博导)曾 定 教 授郑文竹 教 授

2003年10月27日



博 士 学 位 论 文

戊型肝炎病毒衣壳蛋白体外组装相关结构域的研究

Study of assembly associated domain of hepatitis E
virus capsid protein *in vitro*

李 少 伟

指导教师：陈睦传 教授、夏宁邵 研究员



2003年10月27日

厦门大学原创性说明

本人郑重声明：所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除了文中已注明引用的内容外，本论文不含任何其它个人或集体已经发表或撰写过的作品成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式表明。本人完全意识到本声明的法律结果由本人承担。

学位论文作者签名：

日期： 年 月 日

目 录

中文摘要	1
英文摘要 (Abstract)	4
缩 略 词 (Abbreviations)	7
前 言	9
1. 蛋白质-蛋白质相互作用的机理	9
1.1. 蛋白质-蛋白质相互作用的结构与组成	9
1.2. 蛋白质-蛋白质相互作用的影响因素	10
1.3. 蛋白质-蛋白质相互作用的构象变化	12
1.4. 蛋白质-蛋白质相互作用的热力学和动力学	12
1.5. 蛋白质-蛋白质相互作用的研究方法	14
2. 蛋白质复性的机理	19
2.1. 蛋白质折叠的途径	19
2.2. 蛋白质折叠的能量图和漏斗模型	20
2.3. 蛋白质复性的方法	22
3. 二十面体立体对称的病毒结构及组装机理	22
3.1. “近似等积” (quasi-equivalent) 二十面体病毒颗粒的结构与组装机理	22
3.2. 病毒结构的研究方法	28
3.3. 杯状病毒 Norwalk virus (NV) 颗粒结构的组装	31
4. 戊型肝炎病毒的概况及其结构的研究进展	34
4.1. 戊型肝炎病毒的概况	34
4.2. HEV 的病毒及类病毒结构的研究进展	38
5. 本论文研究的思路、目的和意义	43
材料与amp;方法	45
1. 材 料	45
1.1. 主要仪器	45
1.2. 主要试剂与材料	46
1.3. 常用溶液及培养基配制	46
2. 方 法	49
2.1. 基因克隆	49
2.2. 重组蛋白的表达与纯化	50
2.3. 重组蛋白的性质分析方法	55
2.4. 重组蛋白的结构分析方法	56
2.5. 计算机辅助设计与分析	57
结果与分析	58

第一部分 原核表达的 ORF2 二聚体和六聚体的形成及其相关结构域的研究	58
1. NE2 重组蛋白的聚合状态——二聚体和六聚体	58
1.1. NE2 聚体形式的 SDS-PAGE 和蛋白印迹实验	58
1.2. NE2 多聚体的分子筛层析分析	59
1.3. NE2 多聚体的 MALDI-TOF-MS 分析	62
1.4. NE2 多聚体的水化分子动力学半径(R_H)的测定	63
2. 二聚体和六聚体的聚合相关结构域的研究	64
2.1. 二聚体形成的核心疏水区及其相互作用结构域	64
2.2. 环境氨基酸对聚体构象的影响	76
2.3. 三对二聚体间相互作用形成六聚体的结构域模型	79
2.4. NE2 及其突变多肽的二级结构研究	81
第二部分 原核表达的 HEV 类病毒颗粒的体外组装及其相关结构域的研究	85
1. VLP 体外组装的影响因素及颗粒形态大小	85
1.1. VLP 体外组装的影响因素	85
1.2. HEV 239 类病毒颗粒的形态和大小分析	92
2. HEV 239 VLP 组装相关结构域的研究	98
2.1. VLP 体外组装的一级结构范围	98
2.2. HEV 239 蛋白的二级结构预测	108
2.3. HEV 239 VLP 组装的功能结构域的研究	110
讨 论	119
1. HEV 衣壳蛋白基本结构亚单位——二聚体是由单体间的疏水作用形成的	119
2. 二聚体与主要天然中和免疫表位形成的关系	120
3. 二聚体可进一步形成衣壳子粒结构——六聚体	122
4. 受调控的子粒结构可组装成 VLP	124
5. 化学交联实验是研究蛋白质结构域的重要工具	126
6. VLP 组装结构域与分子开关的初步证据	129
7. 结构域保守的生物学意义	129
8. HEV ORF2 的功能结构域分布	129
9. HEV 239 VLP 的可能组装机理	131
10. HEV 类病毒颗粒体外组装与体内组装的差异	132
小结与展望	134
参考文献	136
致 谢	150
附 录	151
一、在校期间发表论文	151
二、在校期间获得奖励	153

戊型肝炎病毒衣壳蛋白体外组装相关结构域的研究

摘要

蛋白质-蛋白质相互作用主要是通过作用结构域来实现的,结构域是研究蛋白质结构和功能关系的桥梁,相互作用结构域的研究是当前功能蛋白质组学的研究热点。病毒颗粒的组装是蛋白质-蛋白质相互作用的典型,类病毒颗粒(VLP)结构的解析和组装机理的研究对探讨病毒的感染、发病和致病机理有着重要的理论和实际意义。本论文研究主要利用原核系统表达 HEV 衣壳蛋白,获得衣壳组装的中间态复合体和体外组装的 HEV 类病毒颗粒,并借助多种实验技术对 HEV 衣壳蛋白的体外折叠、相互作用和 VLP 体外组装的相关结构域进行研究,得到了初步的 HEV ORF2 衣壳蛋白的功能结构域分布图,并提出 HEV 类病毒颗粒可能的体外组装机理。

SDS-PAGE 和蛋白印迹实验(WB)的结果显示,HEV ORF2 aa394-606(NE2)多肽的单体间可发生相互作用形成二聚体,三对二聚体还可进一步相互作用形成三邻体(六聚体);高效分子筛层析(HPSEC)和动态光散射(DLS)分析结果均显示,NE2 在溶液中主要以六聚体的形式存在;基质辅助激光解吸电离飞行时间质谱(MALDI-TOF-MS)测得 NE2 单体分子量为 22,899.18Da,六聚体的分子量为 139,705.91Da。根据已有的 HEV 类病毒颗粒三维结构,NE2 二聚体和六聚体可能分别模拟了 HEV 衣壳的基本结构亚单位和子粒结构单位两种衣壳组装的中间态复合物。

为了探讨二聚体形成的关键区域,以及聚体形成与主要天然中和免疫表位的关系,本论文通过末端缺失、延伸和定点突变技术研究 NE2 蛋白的聚合现象。(1) C 端缺失实验结果发现 NE2 蛋白 C 端缺失至少位于 aa601,所得多肽无法从尿素变性液中折叠复性;进一步的 601 位氨基酸定点突变实验显示,替代氨基酸含强亲水性侧链(亲水性指数小于 $-34\text{kJ}\cdot\text{mol}^{-1}$)不仅破坏了二聚体的形成,而且极大降低了蛋白的折叠复性率,而亲水性弱的替代氨基酸大多不影响蛋白的复性和二聚体的形成;将 aa601 周围的氨基酸依次突变成为强亲水性氨基酸,结果发现 aa597-602 疏水区是 NE2 形成二聚体的核心区域。(2) N 端缺失和 C 端延伸实验显示,NE2 蛋白 N 端至少 65 个氨基酸既不影响同源聚合也不直接参与主要的天然中和表位的形成,但可作为环境氨基酸协助中和表位构象的形成,而这种协助作用可被 C 端延伸的 ORF2 末端肽段所代替。提示 HEV 天然中和免疫表位的形成依赖于特定的疏水二聚体构象。

用 PSIPRED 二级结构预测法和圆二色性(CD)光谱对 NE2 及其突变多肽的二级结构进行研究,结果发现可形成聚体的蛋白 α -螺旋的比例高于不能形成

聚体的多肽，而 β -折叠比例则较低；可形成二聚体但不能形成六聚体的多肽的二级结构特征比例处于两类蛋白之间。说明蛋白质的二级结构模式决定了多肽二聚体和六聚体的形成。CD 光谱还显示六聚体多肽的无规卷曲特征谱峰较不发生聚合和仅可形成二聚体的多肽更为显著，提示无规卷曲结构与六聚体的形成相关。

过去的研究中，我们发现 NE2 N 端延伸至 aa368 得到的多肽 HEV 239 可体外组装形成 HEV 类病毒颗粒。为了研究 HEV 239 多肽体外组装的机理，我们采用切向流技术对高纯度的重组蛋白进行体外组装实验。结果发现硫酸铵浓度不同，颗粒的组装效率也不同，提示疏水相互作用与颗粒组装效率密切相关，而硫酸铵对 NE2 多聚体的形成基本没有影响，提示 aa368-394 疏水肽段可作为 VLP 组装的调控元件。另外，组装的温度不仅影响颗粒组装的效率而且与颗粒的大小有关，提示颗粒体外组装过程主要是由蛋白质-蛋白质相互作用体系的焓驱动的。用 HPSEC、DLS、透射电镜 (TEM) 和原子力显微镜 (AFM) 对 HEV 239 类病毒颗粒进行大小和形态分析，发现 HEV 239 呈直径 13.5-23.0nm 的规整的球状颗粒。HEV 239 VLP 表面有明显的凸起和缺刻形态，与天然 HEV 和杆状病毒表达的 VLP 类似。

HEV 239 多肽的末端缺失和延伸实验发现，N 端位于 aa345 以前的多肽不仅不能体外组装成为 VLP，而且不能形成二聚体和六聚体；C 端位于 aa601 以前的肽段可组装成为颗粒，但不形成疏水二聚体，同样也丧失了主要的天然免疫表位。这些结果提示 VLP 体外组装的肽段的一级结构范围是：N 端位于 ORF2 aa345-370，C 端位于 aa601-660。

进一步运用定点突变、化学交联和生物质谱等技术从结构域层次上对 HEV 衣壳蛋白的组装过程的相互作用进行研究，本研究首次发现了 HEV 衣壳蛋白体外折叠、相互作用和颗粒组装的三个相关结构域：

(1) 将 NE2 二聚体形成核心区域的 aa597A 定点突变成为 C 发现，突变多肽仍可形成疏水二聚体，而且二聚体上的 aa597C 可“零距离”交联形成二硫键二聚体，说明二聚体的 aa597C 在空间位置上相接近，处于可生成化学键的距离以内，提示聚合核心区域 aa345-370 所在的结构域为 HEV 衣壳蛋白单体间疏水相互作用的作用结构域。此外，SDS-PAGE 和 WB 实验显示二硫键二聚体 SDS-PAGE 相对分子量大于疏水二聚体，而且丧失了主要的天然免疫表位。由于二硫键的刚性结构可导致蛋白构象的变化，提示该疏水相互作用结构域具有一定的柔性。

(2) 将 HEV 239 多肽 N 端 28 肽高度保守的 aa372L、aa375L 和 aa395L 突变成强亲水性氨基酸 E，结果发现突变多肽不能进行体外组装，而且不能形成二聚体和六聚体。说明突变位点的 L 对于 N 端 28 肽 aa368-395 的 HLH 结构域有

重要的贡献，结构的破坏可导致体外组装调控失常，甚至影响到 C 端疏水相互作用的发生，提示该结构域是 VLP 体外组装的调控结构域。

(3) 以 DTSSP 对 HEV 239 VLP 进行化学交联，结果发现 HEV 239 VLP 可被交联，提示被交联的结构域为颗粒组装的组装结构域。由于 HEV 239 多肽所含 Lys 的数目和位置与 NE2 相同，而同样条件下 NE2 六聚体不被交联，说明 HEV 239 VLP 上的 NE2 肽段发生了变构。进一步的胰酶酶解肽谱发现，化学交联发生在 aa423K。这些结果提示 aa423K 所在的结构域是 HEV 类病毒颗粒的组装结构域。

综合上述的结果，本文获得了初步的 HEV ORF2 衣壳蛋白体外折叠、相互作用、中和表位形成和 VLP 体外组装的功能结构域分布图，并提出 HEV 类病毒颗粒可能的体外组装机理。从而，为重组 HEV 疫苗的分子设计提供理论依据，为 HEV 的发病机理和细胞受体的研究提供线索。

关键词：戊型肝炎病毒 聚体 类病毒颗粒 体外组装 结构域 子粒

Study of assembly associated domain of hepatitis E virus capsid protein *in vitro*

Abstract

Interface domains play a very important role in protein-protein interaction, and build a bridge between protein's structure and function. A hot theme on functional proteomics is investigating into interface domains of proteins. Self-assembly of virion is a paradigm of protein-protein interaction. The unwrapped structures and elucidated assembly mechanisms of some virus-like particles (VLPs) always leads to insights into infection and pathogenesis of the corresponding viruses. In this study, bacterial expressed hepatitis E virus (HEV) capsid proteins, which can self-assembly into oligomers or VLPs *in vitro*, were used to study the domains related to subunit refolding, interacting and VLPs' assembly.

The peptide NE2 locating to aa394-606 of HEV ORF2 was found to interact with one another to form dimers on SDS-PAGE, and three dimers could further interact trimerically to form hexamers which could only be detected by Western blotting (WB) with monoclonal antibodies. The NE2 hexamers were the predominant form in detergent-free solution, which was verified by high-performance size exclusion chromatography (HPSEC) and dynamic light scattering (DLS). The results of matrix-assisted laser desorption/ionization-time of flight-mass spectrometry (MALDI-TOF-MS) showed that the molecular weight of NE2 monomer and hexamer were 22,899.18Da and 139,705.91Da respectively, both similar with the predicted values. These results indicated that the hexamer be another stable assembly intermediate besides dimer, which might mimic the capsomers and basic structural subunits of HEV, respectively.

In order to discover the core domain which contributes to dimerization and if the formation of main neutralizing epitope is associated with the dimerization of the peptide, truncated, extended, and site-substitution mutants of NE2 protein were produced and their dimerization property and the neutralizing epitope activity were tested. The results showed that C-terminal from aa601 truncated mutants could hardly refold from urea solution. Site-substitution of the Leu on aa601 of NE2 by any other amino acids with hydrophilicity value less than $-34\text{kJ}\cdot\text{mol}^{-1}$ resulted in the failure of dimerization and renaturation, while mutants substitution of the same amino acid with low hydrophilic amino acids can dimerize and have the similar neutralizing epitopes' activity as NE2 protein. Substitutions of other amino acids around aa601 with Glu showed that the high hydrophobic region from aa597 to aa602 is the core region

contributing to dimerization of NE2 protein, and also is necessary for the formation of neutralizing epitope. It has no effect on the dimerization when a 65 amino acids fragment of N-terminus of NE2 peptide was truncated. However, the fragment may act as helper for the formation of neutralizing epitope on NE2, because the truncated mutant had lost its reactivity with neutralizing monoclonal antibody, but recovered its reactivity when the C-terminus of this mutant was extended to aa660 of ORF2. These results suggest that the formation of neutralizing epitope depends on the “right” dimerization of the peptides.

Program based secondary structure prediction and the circular dichroism (CD) spectrometry of NE2 and its mutants were performed. The results showed that the secondary structural features of proteins which can dimerize, hexamerize and form the neutralizing epitopes as NE2 are similar, but all distinct from the non-reactive and non-dimerized mutants, and the secondary structural features of mutant which is well-antigenic and can dimerize but cannot hexamerize. It indicated that the formations of dimers and hexamers are related to the secondary structural features of peptides. CDs showed that the random coil structural of hexamer forming peptides were more prominent than others, which suggested that random coil structure benefit the formation of hexamer.

The peptide HEV 239, a N-terminal extension mutant of NE2 protein located on aa368-606 of ORF2, was found earlier to have the property of self assembly into VLPs. Highly-purified denaturated HEV 239 peptide was renatured with a tangential flow device under different conditions. The results showed that the concentration of $(\text{NH}_4)_2\text{SO}_4$ has obviously effect on the assembly of VLPs. Furthermore, the VLPs' size could be varied in different temperature. $(\text{NH}_4)_2\text{SO}_4$ was regarded to effect the hydrophobic status of protein, but it had no effect on the formation of NE2 polymers. Therefore, it was deduced that the N-terminal extension fragment of HEV 239 than NE2, which corresponded to the aa368-394 region of ORF2, might act as a regulator for VLPs' assembly, and the self-assembly *in vitro* of VLPs might be an entropy-driven reaction. The size and morpha of HEV 239 VLPs were analyzed by HPSEC, DLS, transmission electron microscopy (TEM) and atomic force microscopy (AFM). VLPs were visualized as regular spherical particles of ~18nm (13.6-23.0nm) in diameter. There are some indentations and spikes on VLPs' surface, similar to native HEV virions and insect cell expressed VLPs.

It is well-known that primary structure of protein contains all the information of protein folding and specific function. The investigation of HEV 239 peptide by terminal truncation and extension showed that: (1) the peptides, N-terminal of which located from aa345 to aa112, can not self-assembly *in vitro* or form dimers; (2) the peptides, C-terminal of which located upstream to aa601, can self-assembly *in vitro*,

but cannot form dimers and lose their reactivity with neutralizing monoclonal antibody. So the N-terminal amino acids region for peptides which are capable of self-assembly *in vitro* might be located on aa345-370, while their C-terminal region be located on aa601-660. The N-terminal of HEV 239 was predicted to form a Helix-Loop-Helix domain, which is a conserved protein-protein interaction domain on many other proteins. This domain might play important role on regulating VLPs' self-assembly *in vitro*.

With a method combining chemical cross-linking with MALDI-TOF-MS, three domains related to subunit refolding, interacting and self-assembly to VLPs were found:

(1) When the alanine on aa597 was substituted with cysteine, besides the hydrophobic dimer, a disulfide dimer was also found, which indicated that two alanines on aa597 of NE2 dimers are very closely on conformation of the dimers, so that the aa597-602 domain might be the interface domain for hydrophobic dimerization of HEV capsid proteins. Furthermore, the cross-linked disulfide dimers shifted faster than hydrophobic dimers on SDS-PAGE, and its neutralizing epitope was destroyed, which suggested that this interface domain might be flexible and the flexibility is susceptible of rigid structure caused by disulfide bond.

(2) Site-substitution mutants of the aa372L, aa375L or aa395L on the HLH domain of HEV 239 with Glu cannot assembly into VLPs. These sites are also high conserved in wild HEV isolates. These results suggest that the aa368-395 be regulation domain for self-assembly *in vitro* of VLPs.

(3) HEV 239 VLPs could be cross-linked by DTSSP cross-linker, which indicated that the cross-linked domain might be assembly interface domain on VLPs. The result of tryptical peptide mass fingerprinting (PMF) using MALDI-TOF-MS showed that the cross-linking amino acid be the Lys on aa423. However, NE2 peptide which also contains this amino acid cannot be cross-linked with the same condition, which indicated that the conformation structures around this site be different between NE2 and HEV 239. This result suggested that the assembly domain might contain a molecular switch for locally direct binding of intermediates.

Based on all these results, a preliminary profile of the functional domains located on HEV ORF2 was presented and a possible assembly mechanism *in vitro* was proposed, which paved the way for further study of HEV's cellular receptor and anti-HEV vaccine.

Key words: Hepatitis E virus, oligomer, Virus-like particle, self-assembly *in vitro*, domain, capsomer

Abbreviations (缩略词)

- AD: transcription Activating Domain, 转录激活区
- AFM: Atomic Force Microscopy, 原子力显微镜
- Amp: Ampicillin, 氨苄青霉素
- AU: Analytical Ultracentrifugation, 分析型超速离心技术
- BD: DNA Binding Domain, DNA 结合区
- BIA: Biospecific Interaction Analysis, 生物特异相互作用分析
- CCF: Conformation Compressing Factor, 构象压缩比因子
- CCMV: Cowpea Chlorotic Mottle Virus, 豇豆黄斑病毒
- CD: Circular Dichroism, 圆二色性
- cryo-EM: cryo-Electron Microscopy, 冰冻蚀刻电子显微镜
- Da: Dalton, 道尔顿
- DHB: Dihydroxybenzoic acid, 二羟基苯甲酸
- DLS: Dynamic Light Scattering, 动态光散射
- ELISA: Enzyme-Linked ImmunoSorbant Assay, 酶联免疫吸附测定
- EM: Electron Microscopy, 电子显微镜
- ER: Endoplasmic Reticulum, 内质网
- ET-NANBH: Enterically Transmitted Non-A Non-B Hepatitis, 急性传染性非甲非乙型肝炎
- GSH: reduced glutathion(e), 还原型谷胱甘肽
- GSSG: oxidized glutathion(e), 氧化型谷胱甘肽
- GST: Glutathion S-Transferase, 谷胱甘肽 S-转移酶
- HCCA: α -cyano-4- hydroxycinnamic acid, α -氰基-4-羟基肉桂酸
- HE: Hepatitis E, 戊型肝炎
- HEV: Hepatitis E Virus, 戊型肝炎病毒
- HIC: Hydrophobic Interaction Chromatography, 疏水相互作用层析
- HLH: Helix-Loop-Helix, 螺旋-环-螺旋
- HPLC: High Performance Liquid Chromatography, 高效液相色谱
- HPSEC: High Performance Size Exclusion Chromatography, 高效分子筛层析
- IEM: Immuno-Electron Microscopy, 免疫电镜术
- Iso-PrIME: Isoionic-Preparative Isoelectric Membrane Electrophoresis, 制备型等电聚焦膜电泳技术
- Kan: Kanamycin, 卡那霉素
- kD: kilo Daltons, 千道尔顿
- mAb: monoclonal Antibody, 单克隆抗体

MALDI-TOF-MS: Matrix-Assisted Laser Desorption/Ionization- Time Of Flight- Mass Spectrometry, 基质辅助激光解吸电离飞行时间质谱

MAPK: Mitogen-Activated Protein Kinase, 有丝分裂原激活的蛋白激酶

MD: Molecular Dynamics, 分子动力学

MW: Molecular Weight, 分子量

NCBI: National Center for Biotechnology Information, 美国国家生物技术信息中心

NMR: Nuclear Magnetic Resonance, 核磁共振

NV: Norwalk Virus, 诺沃克病毒

ORF: Open Reading Frame, 开放读码框架

PDB: Protein Data Bank, 蛋白质数据库

PhMV: Physalis Mottle Virus, 空泡斑点病毒

PMF: Peptide Mass Fingerprinting, 肽指纹图谱

R_H: Hydrated molecular dynamic Radius, 水化分子动力学半径

RM: Resonant Mirror, 镜面谐振

Rs: Stokes' Radius, 斯托克斯分子半径

RT: Remained Time, 保留时间

SA: Sinapinic Acid, 芥子酸

SBMV: Southern Bean Mosaic Virus, 南豆花叶病毒

SE: Sedementation Equilibrium, 沉降平衡

SPR: Surface Plasmon Resonance, 表面等离子共振

SMV: Sesbania Mosaic Virus, 田菁花叶病毒

SRP: Signal Recognition Protein, 信号肽识别蛋白

SV: Sedimentation Velocity, 沉降速度

T: Triangulation number, 三角划分数即 T 值

TBSV: Tomato Bushy Stunt Virus, 马铃薯矮小病毒

TCV: Turnip Crinkle Virus, 芜菁甘蓝皱纹病毒

TEM: Transmission Electron Microscopy, 透射电子显微镜

TMV: Tobacco Mosaic Virus, 烟草花叶病毒

TYMV: Turnip Yellow Mosaic Virus, 芜菁甘蓝黄叶病毒

Ve: Elution volume of the peotein(s) of interest, 洗脱体积

Vo: Column void volume, 层析柱的外水体积

Vt: Total bed volume, 层析柱的柱床体积

VLP(s): Virus-Like Particle(s), 类病毒颗粒

WB: Western Blotting, 蛋白印迹实验

前 言

天然的蛋白质具有特征的三维空间立体结构，结构决定了蛋白质的功能。结构生物学揭示大量蛋白质分子的精确立体结构及其与复杂生物功能的关系，同时分子遗传学发展了以定位诱变为中心的基因工程技术，为人们改造和研究蛋白质提供工具。然而，蛋白质作为生命的体现者，往往不是单独行使功能的，蛋白质—蛋白质相互作用在细胞的生命过程中扮演着重要的角色。对于蛋白质—蛋白质相互作用的研究，病毒颗粒衣壳立体对称结构的组装及病毒与宿主细胞的相互作用无疑是该领域的典例。病毒作为一种简约的寄生生命形式，以其快速的进化适应环境和逃避宿主的防御体系，对其它生物的生存提出挑战；包括艾滋病毒(HIV)在内的多种动物病毒仍然使人类望而却步、不知所措。从蛋白质—蛋白质相互作用机理、蛋白质折叠机理和体外组装等方面深入研究病毒结构及其组装机理的分子病毒学，必定能够为病毒感染的预防与病毒性疾病的治疗提供思路 and 手段。

1. 蛋白质—蛋白质相互作用的机理

在生物有机体的生命活动中，蛋白质—蛋白质相互作用普遍存在，如免疫系统中的抗原—抗体、生化反应中的酶—底物、信号传导中的效应分子—受体、病毒感染中的病毒—细胞受体以及病毒组装过程中的单体—单体、子粒—子粒等等，均是通过蛋白质—蛋白质的相互作用发挥其特有的生物学功能。蛋白质—蛋白质相互作用一般具有很高的特异性，其中还可能诱发变构现象、级联效应。相互作用过程也是热力学和动力学的过程，伴随着熵、焓和自由能的变化。相互作用的发生取决于蛋白质的结构，目前在 PDB 数据库存有 22,810 个（截至 2003 年 10 月 7 日）的蛋白质结构，展示了形形色色的蛋白质—蛋白质相互作用。

1.1. 蛋白质—蛋白质相互作用的结构与组成

据统计，蛋白质—蛋白质相互作用形成复合物的交界面积约占蛋白单体表面积 的 6-30%，大小为 550-4900Å²，平均面积为 800Å²。在交界表面上，频繁出现的氨基酸多为疏水性氨基酸、丝氨酸、酪氨酸、色氨酸、天冬酰胺、组氨酸和精氨酸；不同的二级结构类型出现也有一定的规律性：无规卷曲(47%)>螺旋(36%)>折叠(17%)^[1,2]。最近，Ma 等^[3]利用多结构比对程序(MUSTA)对 PDB 中具有代表性的分属 10 个蛋白家族的 86 个蛋白复合物进行了接触界面和剩余表面氨基酸的性质和保守性的比较，发现其中色氨酸、苯丙氨酸和甲硫氨酸在蛋白复合物紧密接触的位点(hot spot)极为保守，可用于预测蛋白间的结合位点；而在非结合位点的表面上则无明显的氨基酸保守性。抗原—抗体复合物中抗原上的抗

体结合部位富含色氨酸、酪氨酸、丝氨酸和天冬酰胺残基；与其它蛋白复合物明显不同，芳香族氨基酸可成簇存在于抗体上的结合部位^[4]。

许多稳定的结构域经常参与蛋白复合物的形成^[5]，包括：螺旋-环-螺旋(HLH, helix-loop-helix)^[6,7]、SH2(Src-homology-2)^[8,9]、SH3(Src-homology-3)^[8]、PH (pleckstrin homology)^[8]、PDZ^[10]、PDZ2^[11]、TPR(tetratricopeptide repeat)^[12]、SCAN^[13]、BRCT^[14]和 WW^[15]结构域等。蛋白质结构域可定义为：由至少 35~40 个残基构成的可互换并且半独立的功能单位；是最小、最紧密、也最稳定的一个连贯的三维结构；作为一个结构和功能单位,在同一个蛋白质中有时会重复出现；在高等真核细胞中，对应于一个穿梭外显子^[16]。细胞调控系统常通过蛋白相互作用结构域如 SH2、SH3、PH 和 PDZ 结构域等进行组装^[17]。戊型肝炎病毒(HEV)第三开放读码框编码蛋白(pORF3)的脯氨酸富集区可与细胞蛋白的 SH3 结构域发生结合，活化由细胞有丝分裂原激活的蛋白激酶(MAPK)，从而，参与细胞的信号传导过程^[18]。

1.2. 蛋白质-蛋白质相互作用的影响因素

根据相互作用的机制和稳定性等方面的差异，可将蛋白复合物分成两类：一是临时型复合物（非专性、寿命短），另一是永久型复合物（天然的多聚体结构）^[1]。两类复合物的蛋白质-蛋白质相互作用交界处的结构和性质都不相同：临时型复合物界面是单体蛋白质之间特异性配对的作用位点，类似于酶的活性位点^[19]；永久型复合物的形成可以认为是蛋白质折叠的继续，所以界面性质与蛋白质的核心区或相互作用结构域更类似^[20]。按照这种分类，病毒粒子属于天然的永久型蛋白复合物，而病毒粒子与细胞受体或抗体的结合则形成临时型蛋白复合物。影响蛋白质复合物形成的因素主要包括空间互补性、疏水相互作用、静电吸引力和氢键等。

1.2.1. 空间互补性

分析表明，蛋白质单体间接触部位的结构具有很好的空间互补性。互补的程度取决于蛋白相互作用的类型。一般而言，永久型的天然复合物呈现最高的空间互补性；而临时型复合物的空间互补性较低，其中以抗原-抗体复合物的空间互补性最低^[21]。通常，在蛋白质相互作用的交界表面上存在一些洞穴，约占交界面表面积的 10%左右，其中大约 60%为溶剂分子所填充^[22]。

1.2.2. 疏水相互作用

Tsai 等^[23,24]对 362 个晶体结构中的通过非共价结合的蛋白质-蛋白质相互作用的交界面进行了统计分析(以被包埋的非极性区面积来表示疏水效应的强弱)，发现蛋白质间的交界处较其它表面少含极性/带电氨基酸，而富含疏水性氨基酸。

这提示交界处的核心区是通过疏水效应稳定蛋白质之间的相互作用,同时避免在蛋白质表面的其它位置形成疏水性强的疏水块(Patch),因为这样的疏水块将不利于蛋白质以单体的形式存在于溶液中,从而确保蛋白复合物的单一性。与界面以外的蛋白表面的其它位置相反,蛋白质相互作用交界处的疏水区通常形成1-15个不等的疏水块,面积大小为 $200-400\text{\AA}^2$ 不等^[25]。永久型复合物的接触界面通常具有疏水性表面,疏水性与蛋白质的疏水内核并无本质上的差别^[20],说明这类复合物的形成更类似于蛋白质的折叠过程。对蛋白复合物交界面形态分布的统计结果表明:大约1/3有明显的疏水核,其周围为极性基团所围绕成环,环中为水分子所填充;2/3的界面无明显的连续疏水块(包括短小的疏水块、极性基团和亚基间氢键),从而表现为混合亲水性质,水分子填充在其形成的洞穴里^[5]。在少数情况下,蛋白质-蛋白质相互作用也可以亲水表面相接触,如:酶和多肽抑制剂或底物之间的相互作用^[26]。

蛋白复合物接触表面的平均疏水性数值通常用蛋白的内核和表面疏水性数值的均值来表示。一般而言,疏水相互作用对永久型复合物形成的贡献大于对临时型复合物的作用。前者通常以结合的状态存在于水溶液中,界面的疏水相互作用对于稳定复合物的形成更为有利;而临时型复合物一般只是在水溶液中临时性的聚合而成,低能量效应使得在蛋白亚基的表面不会存在较高的疏水性,故而疏水相互作用对聚合的贡献较小。但是,也有一些临时型的膜蛋白复合物(如细胞色素P450 2B4),与水溶性蛋白相比,主要是依赖于其跨膜部分的疏水相互作用而发生聚合的^[5]。

1.2.3. 静电吸引力

静电吸引力是蛋白质发生相互作用的另一主要作用力^[27]。每个相互作用表面上存在0-12带电基团^[28]。由于这些带电基团易于同其它带电基团发生相互作用、并为亲水性氨基酸残基所围绕,因此,容易产生去溶剂化效应^[29]。现代结构生物学观点认为,蛋白质相互作用交界面是一种静电力的互补性,而不是带电基团的电荷之间的互补^[30]。有人提出,静电力能够促使相遇的蛋白质相互识别,从而进一步形成复合物^[31]。此外,静电力某种程度上决定了这些复合物存在的寿命^[32]。

1.2.4. 氢键

蛋白质折叠形成的氢键大约有76%是由氨基酸残基的侧链基团形成的。蛋白质相互作用的交界面平均含有大约10个氢键,即大约平均每 $100-200\text{\AA}^2$ 的面积上可形成1个氢键^[2]。通常,蛋白质相互作用交界处的氢键,与蛋白质自身折叠时形成的氢键相比,参与基团所处的相对位置较为不利;从能量方面考虑,这些氢键的相互作用也比较弱。蛋白质-蛋白质相互作用部位还可与周围的水分子形成氢键,而且比蛋白质相互作用形成的氢键更强^[28]。

Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.

厦门大学博硕士学位论文摘要库