

螺旋组装体的电镜三维重构

张 艳, 孙 飞

中国科学院生物物理研究所, 生物大分子国家重点实验室, 北京 100101

收稿日期: 2013-11-29; 接受日期: 2013-12-05

基金项目: 国家自然科学基金青年基金项目(31100617), “973”计划项目(2011CB910301), 青年“973”计划项目(2014CB910700)

通讯作者: 孙飞, 电话: (010)64888582, 传真: (010)64888376, E-mail: feisun@ibp.ac.cn

摘要: 低温电镜三维重构技术近年来得到飞速发展, 其在分析生物大分子复合物的结构和功能中起着重要作用。低温电镜三维重构技术主要包括电子断层三维重构技术、电子晶体学及单颗粒三维重构技术。由于无需结晶, 且所能达到的分辨率较高, 目前, 单颗粒三维重构技术已经成为电镜三维重构领域的主流技术。在用单颗粒三维重构技术进行结构解析时, 有一类特殊的研究对象, 即螺旋组装体, 如烟草花叶病毒 (tobacco mosaic virus, TMV)、微管和微丝 (microfilament)、艾滋病病毒 1 型 (HIV-1) 衣壳蛋白等, 对这类螺旋组装体的结构进行三维重构不是很容易, 涉及诸多物理和数学的概念。本文重点介绍螺旋组装体的基本重构原理, 包括螺旋组装体的数学描述、衍射指数的推导和物理意义、螺旋点阵和真实螺旋组装体 Fourier 变换间的关系, 以及螺旋对称参数的测定和计算等。在此基础上, 文章综述了螺旋组装体的三维重构算法, 主要包括频率空间的 Fourier-Bessel 重构方法和实空间的迭代单颗粒重构方法。最后, 文章以与细胞极性发生相关的蛋白 Par-3 的 N 端结构域 Par-3 NTD 为实例, 详细介绍了 Par-3 NTD 螺旋组装体的实空间迭代单颗粒重构方法。

关键词: 螺旋组装体; 低温电镜; 三维重构; Fourier-Bessel 变换; 实空间迭代的螺旋三维重构算法

中图分类号: Q71

DOI: 10.3724/SP.J.1260.2013.30159

引 言

螺旋在日常生活中很常见, 比如弹簧、盘旋楼梯等都具有螺旋结构。这种螺旋结构体在生物体中一样普遍存在, 且在行使生命活动的过程中具有重要作用。沃森和克里克 1953 年发现的 DNA 双螺旋结构及蛋白质中广泛存在的基本结构组成部分 α 螺旋, 都是螺旋结构在生命体中的典型代表。组成细胞骨架的纤维是由微管^[1-4]、肌动蛋白^[5-7]、肌凝蛋白^[8-10] 等球状蛋白按照螺旋点阵组装的。类似的, 组成细胞外纤维的鞭毛^[11]和菌毛^[12,13], 以及细胞外基质的胶原蛋白^[14]等, 也是由组成蛋白以螺旋结构的形式组装而成的。有些病毒, 比如艾滋病病毒 1 型衣壳蛋白也可以组装成螺旋结构^[15,16]。这些例子表明了螺旋组装体的一个重要的共同性质: 由单个的基本构造单元组装成一个可延长的、纤维状的结构。这种纤维结构的性质则由螺旋体的空间几何性质及相邻基本构造单元间的相互作用关系来确定^[17,18]。

近几十年来, 已经有大量的相关研究揭示了螺旋组装体的几何结构, 并发展了相应的算法以解析螺旋组装体的结构。其中最主要的算法可以分为两类: 一类是基于 Fourier-Bessel 变换的 Fourier-Bessel 重构^[17,19], 该方法主要基于 Cochran 等 1952 年对螺旋组装体衍射性质的研究^[20,21], 根据 Fourier 空间衍射信息, 确定相应 Bessel 函数阶数的衍射值, 然后再进行逆 Fourier 变换, 得到实空间的螺旋组装体的结构^[22~25]; 另一类是基于单颗粒三维重构的实空间迭代的螺旋三维重构算法 (iterative helical real space reconstruction, IHRSR)^[26,27]。尤其是 IHRSR 算法发展以后, 关于螺旋组装体的结构研究层出不穷 (如 F-actin^[6,28]、Dynammin^[29]、EspA^[30]等等)。如此众多的螺旋组装体的结构研究, 都离不开螺旋组装体其复杂的重构原理。本文将总结螺旋组装体的基本重构原理, 介绍螺旋组装体的基本数学描述^[31,32]、螺旋结构体与二维点阵的关系、螺旋组装体与二维点阵的 Fourier 变换、衍射信息的提取及其螺旋参数的计算。在这些基本理论知识的基础上, 进一步介绍螺旋组装体的两种主要重构方法的主要步骤。并通过 Par-3 的 N 端结构域的冷冻电镜三维重构, 展示实空间螺旋三维重构的一个具体应用。

螺旋组装体的数学描述及其相关理论基础

螺旋组装体的数学描述

一个连续的螺旋线可以通过柱坐标系以下面的方程组很好地表示出来:

$$\begin{cases} x=r\cos(2\pi z/P) \\ y=r\sin(2\pi z/P) \\ z=z \end{cases} \quad (1)$$

其中, r 表示半径, ϕ 为从 z 轴正方向看自 x 轴按逆时针方向转所转过的角度。方程(1)等价于下面的表达式:

$$\begin{cases} r=r \\ \phi=2\pi z/P \\ z=z \end{cases} \quad (2)$$

其中, P 表示螺距。

以上是螺旋线的连续形式表达式。而生物组装体实际上是连续螺旋结构的离散化 (见图 1A), 即由基本生成元 (亚基) 沿着螺旋轨迹以相等的间距排布。该排布可以看成单个亚基的两种运动方式的合成: 即沿着 z 轴等间距 (Δz) 地上升, 同时又绕着 z 轴等角度 ($\Delta\phi$) 地旋转 (见图 1B); 或者等价地通过螺旋组装体一圈 (即单个亚基绕螺旋轴旋转 360°) 中基本单位的数目来描述。若把基本生成元看作一个抽象的点, 离散的螺旋组装体则是一个在空间中按照旋转角度 ($\Delta\phi$) 和上升间距 (Δz) 两个参数规则排列的螺旋点阵。本文要讨论的正是离散化的螺旋组装体的基本原理, 以及实际生物样品的螺旋三维重构。图 1 是一个抽象的螺旋点阵示意图。图 1A 中一圈所包含基本单位的个数为 5。但对于实际的样品, 一圈中所含基本单位的数目并不一定是整数, 也可能是分数。每个基本单位在 z 方向上的重复距离 c 可以通过表达式 $c=tP-\mu\Delta z$ 来描述, 其中, t 表示一个重复距离 c 中包含的圈数, μ 表示一个 c 中包含的基本单位的个数, P 是螺距, Δz 是相邻两个基本单位间上升的距离。比

如烟草花叶病毒 3 圈 49 个亚基是一个重复，重复距离 c 是 69 \AA 。其中，烟草花叶病毒的单个亚基上升的距离是 1.408 \AA ，两个亚基间的旋转角度 $\Delta\theta$ 是 22.04° ^[33,34]。

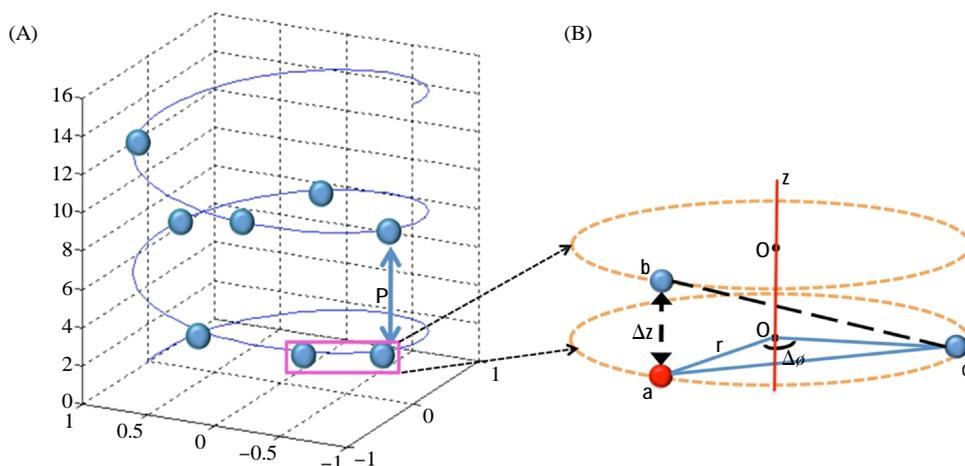


图 1 螺旋结构的几何结构示意图 (A) 螺旋组装体的离散结构示意图。每个蓝色球代表一个基本不对称单位，蓝色曲线代表形成螺旋结构所经过的连续的螺旋线轨迹。P 表示一个螺距的距离；(B) r 表示基本单位离螺旋轴 z 的半径， $\Delta\theta$ 表示从 z 轴正方向看，基本单位围绕 z 轴按逆时针所旋转的角度， Δz 表示基本单位沿着 z 轴方向上升的距离。该图所模拟的螺旋结构一圈中恰好有 5 个基本单位。 z 轴表示螺旋轴，这在螺旋组装体领域的研究中是约定俗成的

Fig.1 The geometry diagram of helical assembly (A) The discrete geometry diagram of helical assembly. Every blue ball is the basic asymmetric unit, the blue curve means the motion path of the blue ball. P represents the distance of one pitch. (B) r denotes the radius; $\Delta\theta$ is the rotation angle around the z axis according to clockwise of positive direction of the z axis. Δz is the translation distance along the z axis. There are exactly 5 subunits among one pitch for this modeling helical assembly. The z axis is usually the helical axis in the helical assemblies structure study field

螺旋点阵与二维点阵的关系

理解二维点阵与螺旋点阵的关系，将有利于理解螺旋组装体的 Fourier 变换及其参数确定。

将一个螺旋点阵沿着其所在柱面的其中一根母线剪开，再平铺到平面上，得到的是一个二维的点阵；同样，在二维点阵中选取两个基准点（该两点之间的连线称为圆周向量）；然后过两个基准点，分别画垂直于圆周向量的垂线；再以两条垂线为边，将两个基准点和两条边重合，得到的柱体即是一个螺旋点阵（图 2），圆周向量正是该螺旋点阵中垂直于轴线的圆周^[7]。在二维点阵中选取不同的基准点，将得到不同的螺旋点阵（图 2）。二维点阵和圆周向量唯一确定了螺旋点阵^[35]。由于螺旋点阵的半径是有限的，螺旋点阵与圆周向量只能相交有限次，该相交次数定义为螺旋点阵的起始数。二维点阵是由两个基向量生成的，而根据选取的基向量的不同，每个螺旋点阵又可看成不同起始数的螺旋线族。比如图 2A 中的黑色线作为螺旋组装体的圆周向量，可以看成不同起始数的螺旋线族。以图中 $(0,0)$ 点为原点， $(0,1)$ 和 $(1,0)$ 分别代表两个基向量方向。取 $(0,1)$ 方向为螺旋点阵的走向，则对应黑色圆周向量的螺旋点阵（图 2B）对应着起始数为 7 的螺旋线族；而若取 $(0,1)$ 和 $(-1,2)$ 方向为基向量，螺旋点阵沿 $(-1,2)$ 方向与黑色圆周向量相交 14 次，则图 2B 可以

看成起始数为 14 的螺旋族。最小的起始数所对应的螺旋线族称为螺旋组装体的基本螺旋线族。

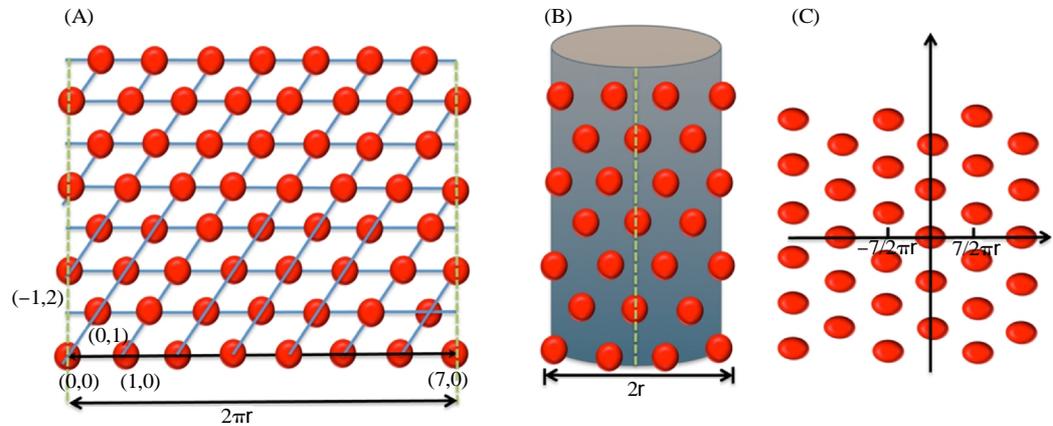


图 2 二维点阵与螺旋点阵的关系 (A) 二维点阵示意图。黑色的双箭头线表示选取的圆周向量, 绿色虚线垂直于黑色的圆周向量。(B) 螺旋点阵示意图。将黑色箭头所指的两点(0,0)和(7,0)重合, 两条绿色虚线重合, 从而将点阵所在的面卷起, 即得到该螺旋组装体, 虚线方向既是螺旋结构的螺旋轴方向。(C) 二维点阵的 Fourier 变换

Fig.2 The relationship between the plane lattice and the helical lattice (A) The diagram of the plane lattice. The black double arrow denotes the circular-vector. The two green broken lines are vertical to the black circular-vector. (B) The diagram of the helical lattice. The helical assembly is formed by rolling up the plane lattice with the two green broken lines and the two points (0,0) and (7,0) coincided with each other. The direction along the broken green line is the helical axis direction. (C) The Fourier transform of plane lattice

以上描述的实际也是从一个二维点阵生成一个螺旋点阵的过程。所生成的螺旋点阵实际上是单个点按照螺旋对称性沿着柱面运动的过程, 如前所述。除了该螺旋对称性, 一个螺旋点阵根据其最小起始数的个数, 又相应地存在 k 次旋转对称性。即: 最小起始数是几, 就具备几次旋转对称性。如图 2B 所示的螺旋组装体还具备 C_7 对称性。图 2B 实际是从图 2A 的平面点阵中取 (0,0) 点和 (7,0) 点间的向量为圆周向量, 将点阵卷起, 让图 2A 中两条绿线 (与黑色圆周向量垂直) 重叠而得到的。(1,0) 方向有 7 个格点 [也是 (0,1) 方向的螺旋线与黑色圆周向量的交点] 均匀地分布在圆周向量上, 因此, 图 2B 除了螺旋对称性外, 同时具备 7 次旋转对称性 (C_7); 反之, 若螺旋点阵具备 k 次旋转对称性, 那么, 任何一条螺旋线都对应着 k 条相同螺旋参数的螺旋线族。所以, 具有 k 次旋转对称性的螺旋组装体的起始数一定是 k 的倍数, 而 k 同时又是基本螺旋线族的起始数, 或者是该螺旋组装体的最小起始数。比如, 图 2B 的基本螺旋线族是起始数为 7 的螺旋线族, 其所对应的其它类型的螺旋线族的起始数均为 7 的倍数, 如上述提到的起始数 14 的螺旋线族。以上论述说明, 一个螺旋组装体可以看成不同的起始数的螺旋结构。这些起始数的最小公约数对应着螺旋组装体的旋转对称性 C_k 的 k 值。

由图 2 也不难理解, 二维点阵衍射空间中的横坐标, 恰好对应着不同螺旋线族所分割圆周向量的均值的倒数。以起始数为 7 的基本螺旋线族为例, 其分割圆周向量 7 次, 则横向上的等分间距为 $2\pi r/7$ (图 2A)。定义 Fourier 空间中横向坐标为 R , 则该族螺旋线在 Fourier 空间横向的坐标 $R=7/2\pi r$, 即 $2\pi r R=7$ 。但对于螺旋组装体, 由于其衍射信息受到柱面曲率的

影响，其衍射信息在横向上被拉伸，即衍射点由二维点阵对应的衍射点变为了拉伸的线，该拉伸实际上就是 Bessel 函数的调制作用引起的，对应的 Bessel 函数的阶数恰好是该衍射信息所对应的螺旋线族的起始数。

螺旋点阵中还存在一个垂直于螺旋轴的二次旋转轴。由于生物分子的手性，真实的螺旋组装体并不一定存在这个二次轴。DNA 双螺旋结构具有这个垂直于螺旋轴的二次轴，即 DNA 双螺旋结构具有 D1 对称性，其结构的顶端和底端是等价的。而对于不存在这种垂直于螺旋轴的二次轴的螺旋组装体，其两端的结构是不一样的，如 Actin 就是这种螺旋组装体。因此，Actin 也被称为有极性的螺旋组装体，其顶端结构和底端结构具有不同的性质^[36]。

综上所述，螺旋组装体中存在两大类对称性：第一种是沿着螺旋轴方向的旋转和上升，第二种是关于螺旋轴的 C_k 或 D_k 次旋转对称性。

螺旋组装体的 Fourier 变换

目前，针对螺旋组装体的冷冻电镜三维重构，无论是 Fourier 空间重构还是在实空间直接利用单颗粒方法进行螺旋重构，都离不开螺旋组装体及其二维投影图的 Fourier 变换。同时，螺旋组装体的三维重构中，螺旋参数的确定方法通常都是在 Fourier 空间确定的。因此，清楚地理解螺旋组装体相关的 Fourier 变换，对于理解螺旋组装体参数间的相互关系，以及利用 Fourier 变换的方法进行螺旋参数的标定，是非常重要的。

柱坐标系下的 Fourier 变换

函数 $f(\mathbf{x})=f(x,y,z)$ 的 Fourier 变换是： $F(\mathbf{X})=F(X,Y,Z)$ ，这里

$$F(\mathbf{X}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\mathbf{x}) \exp(2\pi i \mathbf{X} \cdot \mathbf{x}) d^3\mathbf{x} \quad (3)$$

$F(\mathbf{X})$ 的逆变换是：

$$f(\mathbf{x}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(\mathbf{X}) \exp(-2\pi i \mathbf{x} \cdot \mathbf{X}) d^3\mathbf{X} \quad (4)$$

实空间和 Fourier 空间的柱坐标系分别定义如下：

$$\begin{cases} x=r\cos\phi \\ y=r\sin\phi \\ z=z \end{cases} \quad (5)$$

$$\begin{cases} X=R\cos\Phi \\ Y=R\sin\Phi \\ Z=Z \end{cases} \quad (6)$$

由式(3)、(4)、(5)和(6)，可推得柱坐标系下螺旋组装体的 Fourier 变换是：

$$F(\mathbf{R},\Phi,Z) = \int_{-\infty}^{\infty} \int_0^{2\pi} \int_0^{\infty} f(r,\phi,z) e^{2\pi i [R\cos(\Phi-\phi)+Zz]} r dr d\phi dz \quad (7)$$

前面已经提到，螺旋组装体关于 ϕ 和 z 具有周期性。在二维结构中，如果在 x 和 y 方向有周期性（比如二维晶体），其 Fourier 变换在其倒易空间呈现层次分布（见图 3）。

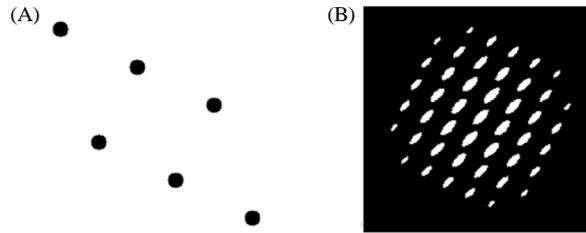


图 3 (A) 模拟的规则排列的二维点阵; (B) 图 A 的 Fourier 变换, 呈现规则的层状排列

Fig.3 (A) Modeling plane lattice with regular array; (B) The Fourier transform of Fig.A with regular layer array

这个性质也可以利用螺旋结构组装体关于 ϕ 和 z 的周期性推广到螺旋结构的 Fourier 变换中去。二维的周期性结构的 Fourier 变换中含有 $\exp[2\pi i(Xx+Yy)]$ 。因此, 我们期望式(7)也有一个类似的简化。 Z 和 z 在式(7)中已经有类似的表达, 我们只需对 Φ 和 ϕ 做以下变换, 以得到简化的形式:

利用欧拉公式 $e^{i\phi} = \cos(\phi) + i\sin(\phi)$, 可得

$$\sin(\phi) = (1/2i)(e^{i\phi} - e^{-i\phi}) \tag{8}$$

由 $\cos(\Phi - \phi) = \sin(\Phi - \phi + \pi/2)$ 及式(8), 容易推出:

$$e^{2\pi i(rR\cos(\Phi - \phi))} = e^{2\pi i(rR(1/2)(u - u^{-1}))} \tag{9}$$

其中, $u = e^{i(\Phi - \phi + \pi/2)}$ 。

第一类 Bessel 函数的母函数 $e^{(1/2)x(t-t^{-1})}$ 的级数展开^[37]为:

$$e^{(1/2)x(t-t^{-1})} = \sum_{n=-\infty}^{\infty} J_n(x) t^n \tag{10}$$

其中, $J_n(x)$ 是第一类的 Bessel 函数, n 代表 Bessel 函数的阶。由该展开式, 式(7)中的指数部分变为:

$$e^{2\pi i(rR(1/2)(u - u^{-1}))} e^{2\pi i z Z} = \sum_{n=-\infty}^{\infty} J_n(2\pi r R) e^{in(\Phi - \phi + \pi/2)} e^{2\pi i z Z}$$

将上式代入式(7)得到:

$$\begin{aligned} F(R, \Phi, Z) &= \int_{-\infty}^{\infty} \int_0^{2\pi} \int_0^{\infty} f(r, \phi, z) \sum_{n=-\infty}^{\infty} J_n(2\pi r R) e^{in(\Phi - \phi + \pi/2)} e^{2\pi i z Z} r dr d\phi dz \\ &= \sum_{n=-\infty}^{\infty} e^{in(\Phi + \pi/2)} \int_{-\infty}^{\infty} \int_0^{2\pi} \int_0^{\infty} f(r, \phi, z) J_n(2\pi r R) e^{-in\phi} e^{2\pi i z Z} r dr d\phi dz \end{aligned} \tag{11}$$

通常, 为了简化上式, 定义

$$g_n(r, Z) = \int_0^{2\pi} \int_{-\infty}^{\infty} f(r, \phi, z) e^{-in\phi} e^{2\pi i z Z} d\phi dz \tag{12}$$

从而式(11)变为:

$$F(R, \Phi, Z) = \sum_{n=-\infty}^{\infty} e^{in(\Phi + \pi/2)} \int_0^{\infty} g_n(r, Z) J_n(2\pi r R) r dr$$

再定义

$$\begin{aligned} G_n(R, Z) &= \int_0^{\infty} g_n(r, Z) J_n(2\pi r R) r dr \\ &= \int_{-\infty}^{\infty} \int_0^{2\pi} \int_0^{\infty} f(r, \phi, z) J_n(2\pi r R) e^{-in\phi} e^{2\pi i z Z} r dr d\phi dz \end{aligned} \tag{13}$$

则螺旋组装体的 Fourier 变换可以简化为：

$$F(R, \Phi, Z) = \sum_{n=-\infty}^{\infty} e^{in(\Phi+\pi/2)} G_n(R, Z) \quad (14)$$

式(12)和(13)正是螺旋三维重构领域常提到的大 G 函数和小 g 函数，它们对应着 Fourier 空间中的不同层线的衍射信息。

利用复指数函数间的正交性，由式(14)可以推出

$$G_n(R, Z) = \frac{1}{2\pi i^n} \int_0^{2\pi} e^{-in\Phi} F(R, \Phi, Z) d\Phi$$

等式(13)将 $G_n(R, Z)$ 表示为 $g_n(r, Z)$ 的 Hankel 变换。利用 Hankel 变换的逆变换，可得：

$$g_n(r, Z) = \int_0^{\infty} G_n(R, Z) J_n(2\pi r R) 2\pi R dR \quad (15)$$

类似上面的推导，可以利用柱坐标系下的逆 Fourier 变换得到：

$$f(r, \theta, z) = \sum_{n=-\infty}^{\infty} e^{in\theta} \int_{-\infty}^{\infty} g_n(r, Z) e^{-2\pi iz Z} dZ \quad (16)$$

螺旋对称性在 Fourier 空间中的作用

前面已经介绍过螺旋组装体的对称性是由 z 方向的上升 Δz 及绕着 z 轴的旋转 $\Delta\theta$ 确定的。用螺旋组装体的密度函数描述螺旋对称性，则有：

$$f(r, \theta, z) = f(r, \theta + \Delta\theta, z + \Delta z) \quad (17)$$

式(17)两边同时 Fourier 变换，左边 Fourier 变换既是式(11)，右边 Fourier 变换则是式(11)的指数项的指数多加一项 $2\pi Z \Delta z - n \Delta\theta$ 。显然，式(17)的两边同时 Fourier 变换后要保持恒等，必然有： $2\pi Z \Delta z - n \Delta\theta = 2\pi m$ ，其中，m 是使得等式成立的某个整数。或者等价地有下式成立：

$$Z \Delta z = m + n(\Delta\theta/2\pi) \quad (18)$$

以上推理表明，螺旋组装体的 Fourier 变换只能在满足式(18)的 Z 和 n 处有值。通常，式(18)被称为螺旋组装体三维重构中的选择公式。利用螺距 P 与两个螺旋对称参数 $\Delta\theta$ 和 Δz 之间的关系 $P = \Delta z(2\pi/\Delta\theta)$ ，可以将式(18)进一步简化为：

$$Z = m/\Delta z + n/P \quad (19)$$

如果螺旋组装体中存在一个真正的重复距离 c，即 $f(r, \theta, z) = f(r, \theta, z + c)$ ，同样利用等式两边的 Fourier 变换及式(11)，可以推得 $2\pi Z c$ 一定是 2π 的整数倍；即存在整数 L，使得

$$Z c = L \quad (20)$$

将此关系式代入式(19)得到

$$L = m(c/\Delta z) + n(c/P)$$

因此，

$$L = um + tn \quad (21)$$

其中， $u = c/\Delta z$ ，代表螺旋组装体一个最小重复距离 c 中所包含的基本不对称单位的个数； $t = c/P$ ，表示一个最小重复距离 c 中所包含的螺旋圈数。实际上，早在 1958 年，Klug 等人^[21]将式(21)定义为选择公式。后来，该定义进一步地推广到更一般的形式，即式(18)。单纯地利用式(21)并不能给我们带来螺旋组装体的参数信息。而式(18)则可以更直观地表明衍射空

间信息和螺旋参数间的关系。因此，通过螺旋组装体的衍射信息，我们可以绘制其对应的 n - Z 图。目前， n - Z 图被广泛应用到螺旋组装体结构解析的参数确定过程当中。

由展开式(10)，Bessel 函数的阶数 n 可以有无限多个。然而，在真实数据的 Fourier 空间所能衍射出来的信息中， n 是有上限的。其上限 n_m 与原始数据的实际分辨率相关，满足下列不等式：

$$n_m < \{2\pi(r/d)-0.9\}/1.1 \tag{22}$$

其中， r 是螺旋组装体的半径， d 是螺旋结构的冷冻电镜原始数据所能达到的分辨率。

如果螺旋组装体同时具有 k 次旋转对称性，则有 $f(r,\theta,z)=f(r,\theta+2\pi/k,z)$ 。类似前面式(18)及式(21)的推导，同样可以结合式(11)得到 $e^{-2\pi ink}=1$ ，即 n 必须是 k 的整数倍。Bessel 函数的最大值对应的自变量 $2\pi rR$ 与 Bessel 函数的阶数 n 之间存在一定的统计关系 (见表 1)。该最大值又对应着衍射空间在特定层处衍射信息最强的位置。而该位置的横向坐标恰对应着二维平面点阵的衍射位置 (图 2C) 经过调制后的位置，即：二维点阵中 $2\pi rR=n$ (见前面 $n=7$ 的例子)，螺旋组装体的衍射空间中对应层上亮度最强处的变量 R 与 Bessel 函数阶数的关系为： $2\pi rR=1.03n+1$ 或者 $2\pi rR=n+2$ (见表 1)。后者的 R 相对于二维点阵的 R 发生了位置的改变。从而可以看出，螺旋组装体中相应衍射位置处的 Bessel 函数的阶数恰好对应着相应螺旋族的起始数。这也从数学推导的角度证明了前述的结论：具有 k 次旋转对称性的螺旋组装体，其起始数一定是 k 的倍数。同时，当 n 为正时，对应着右手螺旋； n 为负数时，对应着左手螺旋。

表 1 Bessel 函数 $J_n(x)$ 的阶数与最大值处坐标的对应表

Table 1 Table for the orders of the Bessel function $J_n(x)$ which corresponding to the ordinate values of maximum for each order of the Bessel function

$ n $	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	当 $ n $ 较大时
$ x $	0.1	1.8	3.1	4.2	5.4	6.4	7.5	8.6	9.6	10.7	11.8	12.8	13.9	14.9	16	17	$ n +2$ 或者 $1.03 n +1$

以上的讨论告诉我们，螺旋组装体的 Fourier 空间是呈层面状的，见公式(19)及图 4。进一步地，由公式(11)和(19)可知，在不同的衍射层面上，其相位有如下规律：若 n 是偶数，则关于 Z 轴位置对称的点处的相位相差 0；若 n 是奇数，则关于 Z 轴位置对称的点处的相位相差 π 。

螺旋组装体的 Fourier 变换与二维点阵 Fourier 变换的关系

在实际的螺旋组装体的三维重构研究中，我们所能得到的数据是研究对象的一系列二维投影图。这些二维投影图是螺旋组装体的近侧面 (near side) 与远侧面 (far side) 的投影之和。也正是由于近侧面和远侧面的迭加，使得具有螺旋对称结构的样品的冷冻电镜原始颗粒的 Fourier 变换，除了具备 Fourier 空间中的原点对称性以外，还具有关于 Z 轴的对称性。我们可以通过图 5 模拟的二维点阵及相应的 Fourier 变换来理解这一点。

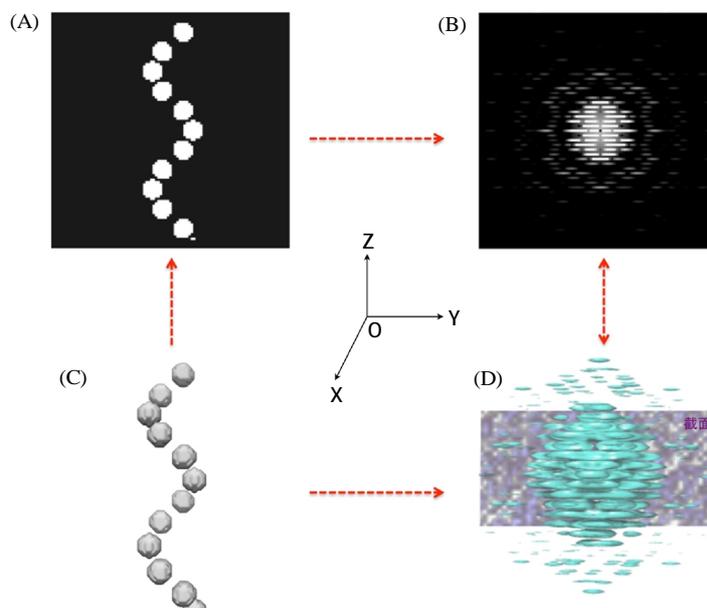


图4 二维投影图及其 Fourier 变换与相应的三维模拟体数据及其 Fourier 变换的关系 (A) 起始数为 1 的螺旋组装体沿着 x 轴方向的投影图; (B) 投影的二维 Fourier 变换; (C) 模拟的起始数为 1 的三维螺旋结构; (D) 三维螺旋组装体的 Fourier 变换。沿着 Z - O - Y 平面的截面即是图 C 沿着 x 轴方向的投影的 Fourier 变换。这里, 实空间的 x - y - z 坐标系与 Fourier 空间的 X - Y - Z 坐标系的坐标轴方向保持一致

Fig.4 The relationship between the plane projection of the three dimension object, the Fourier transform of the plane projection and the modeling three dimension object and its Fourier transform (A) The projection of helical assembly with start number 1. (B) The two dimension Fourier transform of projection. (C) The modeling three dimension helical assembly with start number 1. (D) The Fourier transform of three dimension helical assembly. The cross section along the Z - O - Y plane in the Fourier space is the Fourier transform of projection along the x axis of Fig.C. Here, the axes direction for x - y - z coordinates in real space are consistent with that of X - Y - Z coordinates in Fourier space

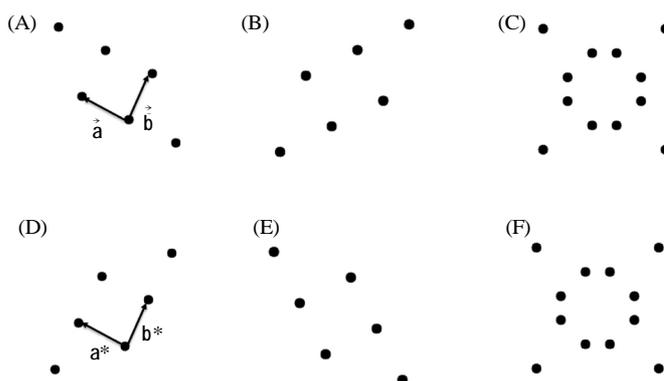


图5 螺旋点阵的 Fourier 变换 (A) 该点阵模拟一个螺旋组装体的远侧面的二维投影点阵; (B) 该点阵模拟的是同一个螺旋组装体的近侧面的二维投影点阵; (C) 该点阵模拟一个螺旋组装体的二维投影, 该点阵恰是点阵 A 和点阵 B 的加和; (D) 点阵 A 的 Fourier 变换; (E) 点阵 B 的 Fourier 变换; (F) 点阵 C 的 Fourier 变换, 该图恰是 Fourier 变换 D 与 E 的加和

Fig.5 Fourier transform of helical lattice (A) The modeling far side projection lattice of helical assembly. (B) The modeling near side projection lattice of helical assembly. (C) The projection of helical assembly with the projection from both far side and near side, C is the sum of A and B. (D) The Fourier transform of Lattice A. (E) The Fourier transform of Lattice B. (F) The Fourier transform of Lattice C, it is exactly the sum of Fourier transform D and E

由中央截面定理知道, 电子束穿过样品后所形成的二维投影图的 Fourier 变换, 恰好是样品的三维 Fourier 空间中穿过中心且与投影方向垂直的截面^[38]。因此, 中央截面定理将螺旋组装体的 Fourier 变换与二维点阵的 Fourier 变换紧密地结合起来^[18,39,40]。有关二维点阵的 Fourier 变换与三维螺旋组装体的 Fourier 变换间的联系, Robert 等合著的《Electron Crystallography of Biological Macromolecules》一书^[38]有更详尽的解释。图 4 展示了二维投影数据及其 Fourier 变换与三维螺旋组装体及其三维 Fourier 变换之间的关系。

Fourier 空间中螺旋参数的推算

有了前面几节的基础, 现在可以讨论如何在 Fourier 空间中利用衍射信息进行螺旋参数的计算。

螺旋组装体垂直于螺旋轴方向的二维投影实际也可以抽象成一系列的正弦波函数信号。起始数为 1 的螺旋组装体的投影对应着一个正弦波函数信号; 起始数为 2 的螺旋组装体的投影对应着两个正弦波函数信号; 以此类推。图 6A 展示了一个正弦波函数信号; 图 6B 是三个正弦波函数信号的迭加, 对应着起始数为 3 的螺旋组装体的投影。

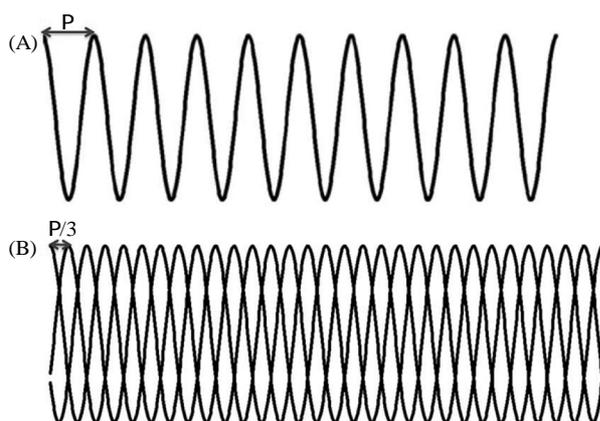


图 6 正弦函数波之间的迭加关系 (A) 一个正弦函数波; (B) 三个正弦函数波的迭加
Fig.6 The superposition relationship of sine waves (A) One sine wave; (B) The sum of three sine waves

实际的生物学样品中的螺旋组装体, 其投影可以看成连续的螺旋组装体的投影 (正弦波函数信号) 与离散的等间距 (相邻不对称单位之间上升的距离 Δz) 排列的直线之间的点积 (图 7A、B、C)。那么, 对应的离散的螺旋组装体的投影的 Fourier 变换则是连续的正弦波函数信号的 Fourier 变换与等间距排列的直线的 Fourier 变换之间的卷积^[29](图 7D、E、F)。

根据倒易空间的关系, 图 7A 中一个螺距是 P , 则其 Fourier 变换中层与层之间的距离是 $1/P$, 也可以通过公式(18)直接推导出来。同理, 图 6B 是起始数为 3 的螺旋组装体的投影, 则其对应的 Fourier 空间中的层线之间的距离是 $3/P$ 。以此类推, 若起始数为 k 的螺旋组装体的螺距是 P , 则其 Fourier 空间层线的高度是 k/P 。同样, 若相邻的不对称单位之间的上升距离是 Δz , 则 Fourier 空间中 Z 轴上 (通常在螺旋重构中将 Fourier 空间的 Z 轴称为子午线) 衍射信息层之间的距离是 $1/\Delta z$ (见图 7C 和 F)。但是实际数据中, 子午线上的这些层线信息受分辨率的约束, 并不一定能看到。通过这段讨论, 我们可以明确地知道在 Fourier 空间中的信息能够反映螺旋组装体的螺距 P , 以及相邻的不对称单位之间的上升距离 Δz 与层线之间的关系。下面将重点讨论如何把这两个参数确定下来。

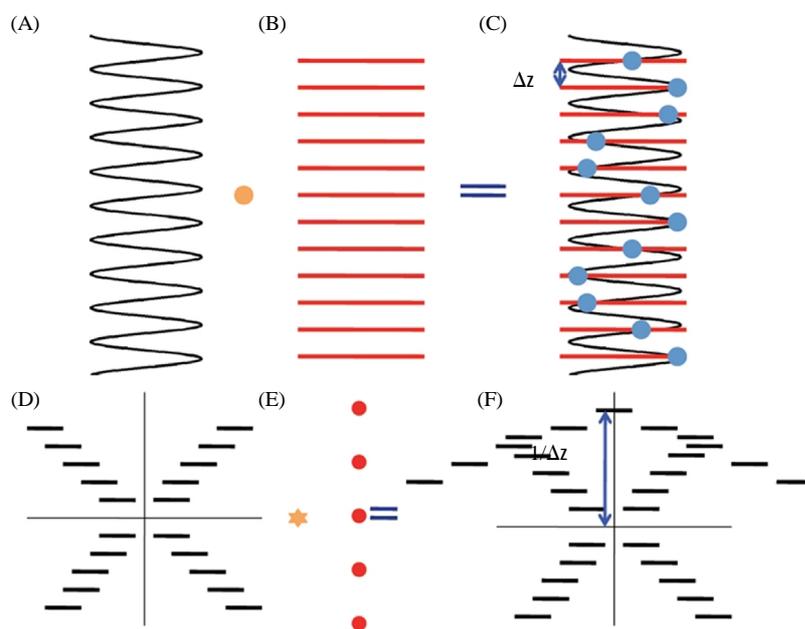


图7 离散螺旋点阵与正弦函数波在实空间及 Fourier 空间的关系 (A) 正弦函数波。(B) 等间距的直线排列。(C) 正弦函数波与等距排列的直线之间的点积。蓝色点阵代表离散的螺旋点阵。(D) 正弦信号波的 Fourier 变换。(E) 等间距的直线排列图的 Fourier 变换。(F) 正弦信号波的 Fourier 变换与等间距直线的 Fourier 变换的卷积

Fig.7 The relationships between discrete helical lattice and sine waves in both real and Fourier space (A) Sine wave. (B) Line array with equal distance of each two neighbour lines. (C) The product of the sine wave and the line array. The blue ball means the discrete helical lattice. (D) The Fourier transform of the sine wave. (E) The Fourier transform of the line array. (F) The Fourier transform of helical lattice (blue balls in Fig.C), which is the convolution of the Fourier transforms of the sine wave and the line array

通常，对于一个生物样品来说，我们事先并不知道它的螺旋参数，得到的只是二维的投影照片。但由二维的投影照片可以得到其衍射空间的信息。因此，充分利用衍射空间的信息成了挖掘螺旋参数的间接途径。如前述，实际的生物样品的二维投影数据可以抽象成有序排列的二维点阵（类似二维晶体）。其基矢向量 \vec{a} 和 \vec{b} 确定了其排布方式（图5A）。在 Fourier 空间，则由基矢向量 \mathbf{a}^* 和 \mathbf{b}^* 确定了其衍射格局（图5D）。若能够通过衍射信息确定出 \mathbf{a}^* 和 \mathbf{b}^* ，即使分辨率不够，也可以利用向量间的线性关系推得 Fourier 空间中子午线上对应 $1/\Delta z$ 处的衍射位置。利用这种思想，在螺旋组装体的三维重构过程中，通常也将收集的原始数据的衍射信息进行基矢向量的标定。根据二维点阵的 Fourier 变换、螺旋组装体二维投影的 Fourier 变换、Bessel 函数的阶数及螺旋组装体的起始数之间的直接联系，通常可通过螺旋组装体二维投影的 Fourier 变换进行 (n, Z) 标定，其中 n 是螺旋组装体中包含的各种可能的螺旋线族的起始数， Z 是螺旋组装体的二维投影的 Fourier 变换的纵向坐标。

若螺旋组装体存在最小重复距离 c ，公式(20)表明，在 Z 方向，不是所有的 Z 都有值，而是在 $Z=l/c$ 处才有衍射信息。因此，在 Fourier 空间，我们也可以重定义坐标向量，通过 (n,l) -plot 来标定衍射方式。但对于螺旋组装体来说，真实的重复距离 c 并不一定存在。且 (n,Z) 标定^[19]较之 (n,l) 标定，能更方便地提供螺旋参数信息。

公式(13)和(19)表明，在 Fourier 空间中，相应变量 R 处的衍射信息与 Bessel 函数 $J_n(2\pi Rr)$ 有关， n 作为 Bessel 函数的阶，同时对应着起始数为 n 的螺旋线族；而起始数为 n 的螺旋线族将圆周向量分割 n 次，每等份的长度是 $2\pi r/n$ (图 2A)；变换到 Fourier 空间，由于螺旋结构本身带有曲率，导致在 Fourier 空间的衍射信息受到了 Bessel 函数的调制，其衍射信息并不像二维点阵 (图 3) 那样对应着一个点，而变成了一条线 (图 4)；该线上衍射信息的强弱与 Bessel 函数的性质 (图 8) 密切相关。由于 Bessel 函数的快速衰减性，函数的最大值处对应着衍射图中最亮的点。而 Bessel 函数取得最大值处的变量 $2\pi Rr$ 与其阶数之间的关系^[19]是： $2\pi Rr=1.03n+1$ 。因此，可以通过 Bessel 函数的这一性质计算 n 值。

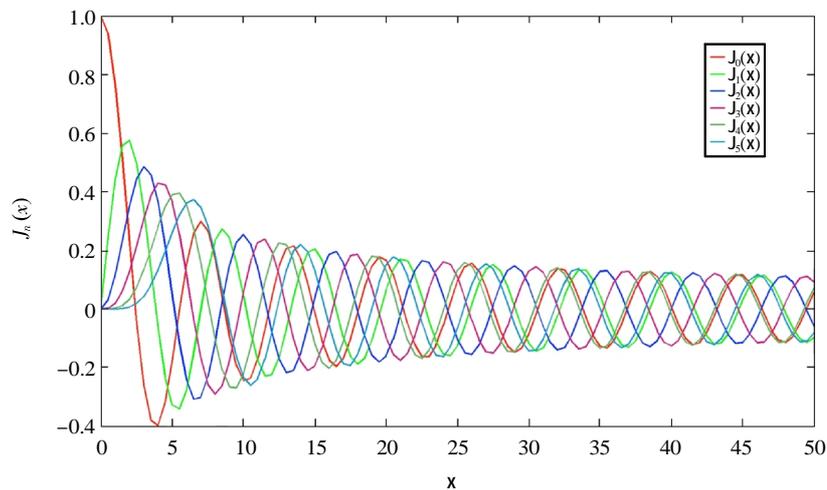


图 8 Bessel 函数曲线图 该图展示了 0~5 阶 Bessel 函数的振幅变化。随着 Bessel 函数阶数的升高，其峰值是向右移动的

Fig.8 The curves according to Bessel functions This figure shows the amplitude variation of Bessel functions with the order 0~5. The position of maximum of Bessel function is moving to right as the order is increasing

如图 9，假设基向量是 (n_1, Z_1) 和 (n_2, Z_2) ，则 n_1 和 n_2 的最大公约数 N 即是螺旋组装体的基本螺旋线族的起始数。根据以上分析，有

$$\Delta z = \frac{N}{|n_2 Z_1 - n_1 Z_2|} \tag{21}$$

$$P = \frac{N}{|k_2 Z_2 - k_1 Z_1|} \tag{22}$$

其中， k_1 和 k_2 可由辗转相除法求解 n_1 和 n_2 的最大公约数 N 时获得，且满足：

$$k_1 n_1 - k_2 n_2 = N \tag{23}$$

有了 Δz 和 P ，容易推得

$$\Delta \theta = 360^\circ \Delta z / P \tag{24}$$

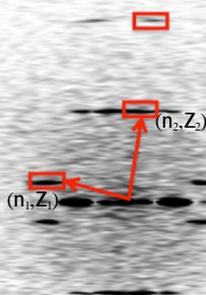


图 9 (n, Z)-plot 的标定 该图实际是 Par-3 N 端结构域的 Fourier 变换，标定出来的向量则是该衍射图对应的 (n, Z) -plot。其中，横坐标对应着 Bessel 函数的阶，纵坐标对应着衍射层线的高度。基本矢量 (n_1, Z_1) 和 (n_2, Z_2) 的选择是任意的，但是一旦选定，必须确保其它衍射层线能够由所选定的基本矢量线性组合而成。相对于二维点阵来说，螺旋组装体投影的 Fourier 变换由于受到 Bessel 函数的调制，半径方向的位置发生了变化。而 Bessel 函数的阶数与 Bessel 函数的最大值在径向的位置之间是一种非线性关系。但是，衍射层在 Z 轴方向的位置是确定的

Fig.9 Index of (n, Z) -plot This is the Fourier transform of the data of Par-3 NTD. The vectors indexed are the (n, Z) -plots according to this Fourier transform. The horizontal ordinate is the order of Bessel function, the vertical ordinate is the Z high value of layer line. The choice of basic vector is arbitrary. But one must make sure that all other diffraction information with maximum intensity in other layer lines can be represented by these two basic vectors. The position of diffraction information in Fourier space of helical assembly is changed due to the modulation of Bessel function. However, it's a nonlinear relationship between the Bessel order and the position where Bessel function with according order has maximum value. But the Z value of each diffraction layer is accurate

目前，也有新的概念引入。Behrmann 等^[42]于 2011 年提出一个新的公约，即用 $(n_1, n_2, \Delta\theta, \Delta z)$ 这个一维向量来统一地定义螺旋组装体系统。这种定义方式的好处在于通过一个一维向量，能够把一个结构体的组装、螺旋对称等信息都囊括进来。但这些概念的引入仍离不开本章基本的数学理论的推导。

Fourier-Bessel 变换方法进行螺旋组装体电镜三维重构

前面重点阐述了螺旋结构的数学基础，在此基础上，我们才能够很好地了解为什么可以利用 Fourier-Bessel 变换的方法进行螺旋组装体的三维重构。下面将主要从具体操作流程上介绍。

第一步：对研究对象所收集的原始照片进行 CTF 矫正，常用的软件有 EMAN1 中的 applyctf^[43]等。

第二步：从研究对象的原始数据集中进行颗粒的挑选。现在有很多程序可以实现螺旋型长条状颗粒的挑选。比如 EMAN1 中的 helixboxer^[43]、XDISPHBOX^[44]、XIMDISP^[45]等。

第三步：进行衍射信息的标定 (如图 9)。根据第一节关于螺旋结构基本数学知识的介

绍, 在衍射信息较好的情况下, 可以直接进行衍射方式的手动标定。目前也可用一些图形界面的软件进行半自动的标定, 比如 WINDEX^[46]、David Stokes 的 EMIP (<http://cryoem.nysbc.org/EmIP.html>) 等。

第四步: 具有相同衍射方式的 CTF 矫正后的颗粒的平均。

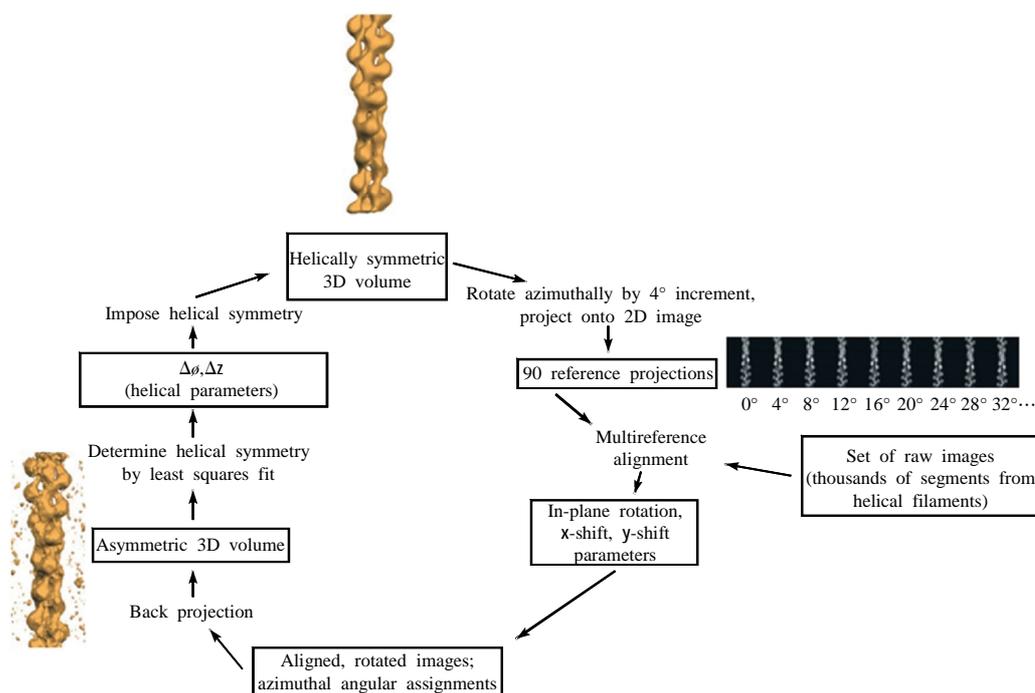
第五步: 三维体数据的计算。有了第四步的平均数据后, 首先进行 Fourier-Bessel 变换, 由公式(15)将 Fourier 空间的层线信息 ($G_n(R,Z)$) 转换到实空间的密度信息 ($g_n(r,Z)$)。然后, 由公式(16)计算出密度函数的值。

以上是用 Fourier-Bessel 变换的方法进行螺旋组装体电镜三维重构的主要步骤。详细的 Fourier-Bessel 重构方法及步骤请参见 Ruben 等的综述文章^[17]。Crowther 等的 MRC^[47]、Whittaker 等的 PHOELIX^[48]、Owen 等的 BRANDEIS^[49]、Yonekura 等的 GUI 程序^[44]、Metlagel 等的 RUBY-HELIX^[50]等软件或软件包, 都包含基于 Fourier-Bessel 方法的螺旋组装体电镜三维重构这一功能。

实空间迭代法进行螺旋结构组装体的电镜三维重构(IHRSR 方法)

前面介绍的 Fourier-Bessel 变换方法在过去的几十年中一直是螺旋组装体电镜三维重构的传统主流方法。但由于螺旋组装体的结构往往不均一, 螺旋参数差一点, 就会组装出完全不同的三维结构; 另外, 有些样品衍射非常弱, 很难得到其 Fourier 空间的衍射信息; 有时候, 同一衍射层所对应的 n 值不唯一, 即所谓的 overlap 现象^[51]。以上也正是在应用基于 Fourier-Bessel 变换的方法进行螺旋结构三维重构时所遇到的困难。尤其在不均一、变形等现象出现时, 不得不对基于 Fourier-Bessel 变换进行螺旋结构三维重构的方法进行改进^[52]。2000 年, Egelman 开发的基于单颗粒三维重构的实空间迭代的螺旋三维重构算法 IHRSR^[26]得到了大家的广泛应用。对于以上提到的 Fourier-Bessel 变换方法的螺旋三维重构所遇到的问题, IHRSR 方法具有一定的优势, 因为它基于单颗粒三维重构的思想, 对于不好的颗粒可以在对位时筛选掉, 所以, 无须对样品的原始数据做取值; 当均一性差、变形等现象出现时, 也可以通过单颗粒三维重构一定程度上进行分类和筛选。但是, 该方法需要先验的螺旋参数, 螺旋参数的确定仍是基于 Fourier-Bessel 变换 [公式(21)、(22)、(24)], 所以, 该方法仍然离不开 Fourier 空间衍射图的标定 (图 9)。图 10 展示了 IHRSR 的重构原理和流程^[26]。2012 年, 该方法已经整合到 SPARX 软件中^[53]。IHRSR 重构算法是在单颗粒三维重构的基础上, 加了一步对三维体数据螺旋参数的强加^[26] (见图 10)。其三维重构的过程使用的是背投影重构。重构时的初始模型通常取为一个圆柱体。2011 年, 孟鑫等以艾滋病病毒 1 型衣壳蛋白的螺旋三维重构为例, 从前期的样品制备到最终的三维重构, 给出了利用 IHRSR 方法进行三维重构的流程^[54]。下面以细胞极性发生相关的蛋白 Par-3 的 N 端结构域为例, 叙述利用 IHRSR 方法进行螺旋组装体三维重构的主要步骤。

Par-3 在调控各类不同物种的细胞极性方面起到基础性的作用。有研究表明 Par-3 在体内可以多聚化。已知 Par-3 蛋白质包含 NTD、PDZ 和 aPKC 结合域等结构域。其中, NTD 结构域的高聚化对于介导 Par-3 的聚集起到不可或缺的重要作用。关于 Par-3 N-端结构域高聚的分子机理, 我们综合利用蛋白质晶体学、低温电镜三维重构技术、原子力显微镜,

图 10 IHRSR 螺旋重构的流程^[26]Fig.10 The flow chart of helical reconstruction method IHRSR^[26]

以及分子动力学模拟等技术方法，深入系统地研究了细胞极化发生因子 Par-3 N-端结构域螺旋自组装的分子机理。研究发现，Par-3 N-端结构域通过静电相互作用形成首尾相接、螺旋上升的组装结构。对相互作用面上的静电相互作用的破坏，使得 Par-3 失去了调控极性的功能^[55]。该结果是利用实空间迭代的螺旋三维重构算法 IHRSR 进行螺旋组装体结构解析得到的。主要步骤如下：

第一步：对研究对象所收集的原始照片进行 CTF 矫正，常用的软件有 EMAN1 中的 applyctf^[43]等。

第二步：从研究对象的原始数据集中进行颗粒的挑选。现在有很多程序可以实现螺旋型长条状颗粒的挑选。比如 EMAN1 中的 helixboxer^[43]、XDISPHBOX^[44]、XIMDISP^[45]等。因为不同直径大小的颗粒一定对应着不同类型的螺旋结构，因此，要将挑选好的颗粒根据直径进行初步的分类。可以将挑选出的颗粒沿螺旋轴向投影，通过测定投影曲线的两个最低峰值之间的距离判定颗粒直径的大小（图 11）。Par-3 的 N 端结构域的直径是 10 nm。

第三步：与基于 Fourier-Bessel 变换的螺旋三维重构一样，对挑选出的颗粒进行 Fourier 变换，然后进行衍射信息的标定，根据衍射图绘制 n-Z 坐标图。根据 n-Z 坐标图，利用公式 (21)、(22)、(24) 把螺旋参数 $\Delta\varphi$ 和 Δz 计算出来。这也是最耗时的一步，在样品不均一的情况下，最好对选取的颗粒逐个进行衍射分析，从而在颗粒直径分类的基础上进一步进行数据的衍射分类。当然，衍射分类也可以先于第二步的直径分类。由于样品厚度的影响，实

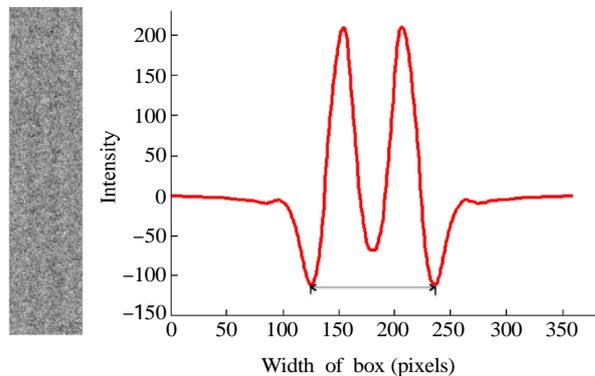


图 11 Par-3 的 N 端结构域的直径测量 左边是原始颗粒，右边是其沿螺旋轴一维投影后的灰度变化曲线。两个最小值之间的距离对应着颗粒的直径

Fig.11 The diameter of Par-3 NTD The left is the raw particle. The right is the intensity variable curve of the one dimensional projection of raw particle. The distance between the two positions according to the two minimum values is the diameter of this particle

实际算出来的两个基本向量的 n 值有一个范围区间。比如，对于 Par-3 的 N 端结构域， $n_1=1$ ， n_2 实际上满足 $6 \leq n_2 \leq 8$ (图 12B)。对于各个可能的 n_2 与 n_1 组合，根据衍射图两个基本向量的相位差判定其奇偶性并进行筛选。Par-3 的 N 端结构域的衍射图对应 $6 \leq n_2 \leq 8$ 的衍射信息，其左右相位差约为 0，所以判定其值必为偶数。所以，最终筛选出来的组合是 $n_1=1$ 、 $n_2=6$ 或者 $n_1=1$ 、 $n_2=8$ 。将这两组组合用来计算各自组合对应的螺旋参数 $\Delta\varphi$ 和 Δz 。

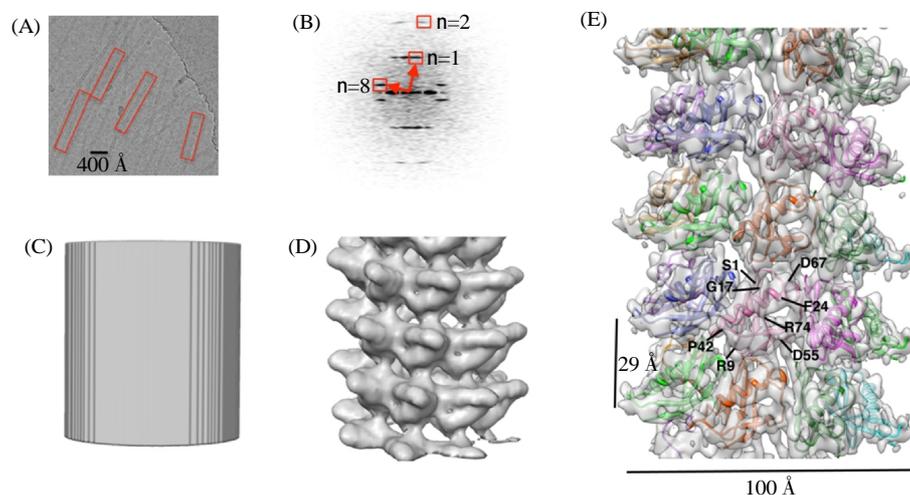


图 12 Par-3 NTD 的冷冻电镜螺旋三维重构 (A) Par-3 NTD 冷冻电镜原始照片。红色区域显示用 EMAN1^[43] 的 heliboxer 所挑选的颗粒。(B) 原始颗粒的衍射图。(C) 模拟的柱体作为 IHRSR 方法的初始模型。(D) IHRSR 方法重构的收敛结果。(E) 用 EMAN1 进一步精修的结果，同时将晶体结构经过分子动力学模拟后匹配到该密度图中

Fig.12 Cryo-electron microscopy helical reconstruction of Par-3 NTD (A) The raw micrograph of Par-3 NTD. The red boxed areas denote the particles selected using heliboxer of EMAN1^[43]. (B) The diffraction pattern of one of the selected particles. (C) The cylinder used as initial model when making reconstruction by IHRSR. (D) The convergent reconstruction result when making reconstruction by IHRSR. (E) The final refinement result of Fig.D by EMAN1. The crystal structure is fitted to the final map using the molecular dynamics simulation

第四步：将上步计算的各组螺旋参数 $\Delta\varphi$ 和 Δz 作为 IHRSR 算法的初始参数，进行实空间的螺旋三维重构。通常，在没有任何先验信息的情况下，模拟一个圆柱作为初始模型（可以利用 SPIDER^[56]、Matlab 等软件）。圆柱作为初始模型，对于初始的投影匹配有一定优势，可以消除某些方向上投影的取向优势等影响。IHRSR 方法的具体迭代次数需要摸索，只要参数最终收敛，即可停止迭代（当然，有时会出现不收敛，这时需要重新检查衍射标定是否正确，同时，数据的质量也会影响投影匹配的效果，进而影响参数的优化，从而导致不收敛的情况发生）。比如，对于 Par-3 的 N 端结构域，最终筛选出的两组参数中， $n_1=1$ 、 $n_2=8$ 这组值对应的螺旋参数 $\Delta\varphi$ 和 Δz 经 IHRSR 方法计算后保持收敛，所以，最终衍射标定确定为 $n_1=1$ 、 $n_2=8$ （图 12B）。

第五步：在追求更高分辨率的情况下，以 IHRSR 算出的结果作为初始模板，进一步利用 EMAN1^[49]等软件进行修正。如图 12D，是用 IHRSR 计算出的收敛后的结构，将作为初始模型进一步进行单颗粒三维重构。在修正时，要考虑到样品的倾斜因素，即由于样品在被快速冷冻时，由于受到表面张力等因素的影响，我们所能收集的原始数据并不一定完全垂直于 z 轴方向，可能会偏离一定的角度范围。所以在修正时，投影角度要包含与 z 轴夹角的一定范围。图 12E 即是最终的修正结果。该图分辨率在 6 Å 左右。

第六步：在有原子坐标的情况下，对修正好的密度图进行分子动力学模拟。图 12E 展示了用 MDFF^[57]和 NAMD2^[58]对 Par3 NTD 进行分子动力学模拟后的匹配效果。若没有原子坐标，可以通过重构结果衍射信息与原始数据衍射信息的比较，以及投影匹配的准确度等辅助手段来验证重构结果的准确性。

讨 论

本文重点讨论了螺旋组装体的数学理论基础，因为这对于衍射方式的标定是尤其重要的。同时，本文就螺旋三维重构的主要步骤进行了阐述。但是，螺旋三维重构的内容还远不止这些。比如，有些螺旋组装体中不对称单位之间有缝隙；有些复杂的螺旋组装体有内、外两层甚至多层，不同层又对应不同的衍射方式；或者螺旋组装体的组装方式动态变化，样品从直径到衍射变化多端，这些情况给螺旋结构的解析带来了更大的挑战^[52]。但是，再复杂的情况，仍离不开本文的这些基础理论做铺垫。同时，在螺旋三维重构领域，难以解决的一个问题就是螺旋结构体手性的确定。通常，通过在数据收集时进行样品的倾转，根据倾转后拍到的电镜照片来识别手性。但数据质量不好的情况下，仍无法直接确定手性，需要进一步的图像技术（比如光学滤镜的传统方法）进行判别^[31]。此外，结合晶体结构匹配程度的好坏，利用原子力显微镜技术等也可以辅助螺旋结构手性的确定。

参考文献:

1. Sandblad L, Busch KE, Tittmann P, Gross H, Brunner D, Hoenger A. The *Schizosaccharomyces pombe* EB1 homolog Mal3p binds and stabilizes the microtubule lattice seam. *Cell*, 2006, 127(7): 1415-1424
2. des Georges A, Katsuki M, Drummond DR, Osei M, Cross RA, Amos LA. Mal3, the *Schizosaccharomyces pombe* homolog of EB1, changes the microtubule lattice. *Nat Struct Mol Biol*, 2008, 15(10): 1102-1108
3. McIntosh JR, Morphew MK, Grissom PM, Gilbert SP, Hoenger A. Lattice structure of cytoplasmic microtubules in a cultured mammalian cell. *J Mol Biol*, 2009, 394(2): 177-182
4. Maurer SP, Fourniol FJ, Bohner G, Moores CA, Surrey T. EBs recognize a nucleotide-dependent structural cap at growing microtubule ends. *Cell*, 2012, 149(2): 371-382
5. Fujii T, Iwane AH, Yanagida T, Namba K. Direct visualization of secondary structures of F-actin by electron cryomicroscopy. *Nature*, 2010, 467(7316): 724-728
6. Galkin VE, Orlova A, Schroder GF, Egelman EH. Structural polymorphism in F-actin. *Nat Struct Mol Biol*, 2010, 17(11): 1318-1323
7. Murakami K, Yasunaga T, Noguchi TQ, Gomibuchi Y, Ngo KX, Uyeda TQ, Wakabayashi T. Structural basis for actin assembly, activation of ATP hydrolysis, and delayed phosphate release. *Cell*, 2010, 143(2): 275-287
8. Woodhead JL, Zhao FQ, Craig R, Egelman EH, Alamo L, Padron R. Atomic model of a myosin filament in the relaxed state. *Nature*, 2005, 436(7054): 1195-1199
9. Alamo L, Wriggers W, Pinto A, Bartoli F, Salazar L, Zhao FQ, Craig R, Padron R. Three-dimensional reconstruction of tarantula myosin filaments suggests how phosphorylation may regulate myosin activity. *J Mol Biol*, 2008, 384(4): 780-797
10. Zhao FQ, Craig R, Woodhead JL. Head-head interaction characterizes the relaxed state of *Limulus* muscle myosin filaments. *J Mol Biol*, 2009, 385(2): 423-431
11. Liu J, Lin T, Botkin DJ, McCrum E, Winkler H, Norris SJ. Intact flagellar motor of *Borrelia burgdorferi* revealed by cryo-electron tomography: Evidence for stator ring curvature and rotor/C-ring assembly flexion. *J Bacteriol*, 2009, 191(16): 5026-5036
12. Mu XQ, Bullitt E. Structure and assembly of P-pili: A protruding hinge region used for assembly of a bacterial adhesion filament. *Proc Natl Acad Sci USA*, 2006, 103(26): 9861-9866
13. Li YF, Poole S, Nishio K, Jang K, Rasulova F, McVeigh A, Savarino SJ, Xia D, Bullitt E. Structure of CFA/I fimbriae from enterotoxigenic *Escherichia coli*. *Proc Natl Acad Sci USA*, 2009, 106(26): 10793-10798
14. LeBleu VS, Macdonald B, Kalluri R. Structure and function of basement membranes. *Exp Biol Med (Maywood)*, 2007, 232(9): 1121-1129
15. Zhang P, Meng X, Zhao G. Tubular crystals and helical arrays: Structural determination of HIV-1 capsid assemblies using iterative helical real-space reconstruction. *Methods Mol Biol*, 2013, 955: 381-399
16. Zhao G, Perilla JR, Yufenyuy EL, Meng X, Chen B, Ning J, Ahn J, Gronenborn AM, Schulten K, Aiken C, Zhang R. Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. *Nature*, 2013, 497(7451): 643-646
17. Diaz R, Rice WJ, Stokes DL. Fourier-Bessel reconstruction of helical assemblies. *Methods Enzymol*, 2010, 482: 131-165
18. DeRosier DJ, Moore PB. Reconstruction of three-dimensional images from electron micrographs of structures with helical symmetry. *J Mol Biol*, 1970, 52(2): 355-369
19. DeRosier D, Stokes DL, Darst SA. Averaging data derived from images of helical structures with different symmetries. *J Mol Biol*, 1999, 289(1): 159-165
20. Cochran W, Crick FHC, Vand V. The structure of synthetic polypeptides I. The transform of atoms on a helix. *Acta Crystallogr*, 1952, 5: 581-586
21. Klug A, Crick FHC, Wyckoff HW. Diffraction by helical structures. *Acta Crystallogr*, 1958, 11: 199-213
22. Amos LA, Klug A. Three-dimensional image reconstructions of the contractile tail of T4 bacteriophage. *J Mol Biol*, 1975, 99(1): 51-64
23. Moody MF. Biophysical electron microscopy: Basic concepts and modern techniques// HPWaV U. Image analysis of electron micrographs. Academic Press, 1990: 170-197, 208-222
24. Smith PR, Aebi U. Studies of the structure of the T4 bacteriophage tail sheath. I. The recovery of three-dimensional structural information from the extended sheath. *J Mol Biol*, 1976, 106(2): 243-271
25. Stewart M. Computer image processing of electron micrographs of biological structures with helical symmetry. *J Electron Microscop Tech*, 1988, 9(4): 325-358
26. Egelman EH. A robust algorithm for the reconstruction of helical filaments using single-particle methods. *Ultramicroscopy*, 2000, 85(4): 225-234
27. Egelman EH. Single-particle reconstruction from EM images of helical filaments. *Curr Opin Struct Biol*, 2007, 17(5): 556-561
28. Galkin VE, Orlova A, Egelman EH. Actin filaments as tension sensors. *Curr Biol*, 2012, 22(3): R96-101
29. Chen YJ, Zhang P, Egelman EH, Hinshaw JE. The stalk region of dynamin drives the constriction of dynamin tubes. *Nat Struct Mol Biol*, 2004, 11(6): 574-575
30. Wang YA, Yu X, Yip C, Strynadka NC, Egelman EH. Structural polymorphism in bacterial EspA filaments revealed by cryo-EM and an improved approach to helical reconstruction. *Structure*, 2006, 14(7): 1189-1196
31. Hawkes PW, Valdre U. Biophysical electron microscopy. London: Academic Press, 1990
32. Glaeser RM, Downing K, DeRosier D, Chiu W, Frank J. Electron crystallography of biological macromolecules.

- Oxford University Press, 2007
33. Kendall A, McDonald M, Stubbs G. Precise determination of the helical repeat of tobacco mosaic virus. *Virology*, 2007, 369(1): 226-227
 34. Clare DK, Orlova EV. 4.6 Å Cryo-EM reconstruction of tobacco mosaic virus from images recorded at 300 keV on a 4kx4k CCD camera. *J Struct Biol*, 2010, 171 (3): 303-308
 35. Baker TS, Caspar DL. Computer image modeling of pentamer packing in polyoma virus "hexamer" tubes. *Ultramicroscopy*, 1984, 13(1-2): 137-151
 36. Murakami K, Yasunaga T, Noguchi TQP, Gomibuchi Y, Ngo KX, Uyeda TQP, Wakabayashi T. Structural basis for actin assembly, activation of ATP hydrolysis, and delayed phosphate release. *Cell*, 2010, 143(2): 275-287
 37. Lebedev NN. Special functions and their applications. Dover, Mineola: NY, 1972
 38. Robert M, Glaeser KD, David DeRosier, Wah Chiu, Joachim Frank. Electron crystallography of helical structures//Electron crystallography of biological macromolecules. Oxford University, 2007: 78-82, 304-342
 39. 张凯, 张艳, 胡仲军, 季刚, 孙飞. 电子显微三维重构技术与前沿. *生物物理学报*, 2010, 26(7): 533-559
Zhang K, Zhang Y, Hu ZJ, Ji G, Sun F. Development and frontier of electron microscopy 3D reconstruction. *Acta Biophys Sin*, 2010, 26(7): 533-559
 40. DeRosier DJ, Klug A. Reconstruction of three dimensional structures from electron micrographs. *Nature*, 1968, 217: 130-134
 41. Toyoshima C. Structure determination of tubular crystals of membrane proteins. I. Indexing of diffraction patterns. *Ultramicroscopy*, 2000, 84(1-2): 1-14
 42. Tsai CJ, Nussinov R. A unified convention for biological assemblies with helical symmetry. *Acta Crystallogr D Biol Crystallogr*, 2011, 67(Pt 8): 716-728
 43. Ludtke SJ, Baldwin PR, Chiu W. EMAN: Semiautomated software for high-resolution single-particle reconstructions. *J Struct Biol*, 1999, 128(1): 82-97
 44. Yonekura K, Toyoshima C, Maki-Yonekura S, Namba K. GUI programs for processing individual images in early stages of helical image reconstruction — For high-resolution structure analysis. *J Struct Biol*, 2003, 144(1-2): 184-194
 45. Smith JM. Ximdisp — A visualization tool to aid structure determination from electron microscope images. *J Struct Biol*, 1999, 125(2-3): 223-228
 46. Ward A, Moody MF, Sheehan B, Milligan RA, Carragher B. Windex: A toolset for indexing helices. *J Struct Biol*, 2003, 144(1-2): 172-183
 47. Crowther RA, Henderson R, Smith JM. MRC image processing programs. *J Struct Biol*, 1996, 116(1): 9-16
 48. Carragher B, Whittaker M, Milligan RA. Helical processing using PHOELIX. *J Struct Biol*, 1996, 116(1): 107-112
 49. Owen CH, Morgan DG, DeRosier DJ. Image analysis of helical objects: The Brandeis helical package. *J Struct Biol*, 1996, 116(1): 167-175
 50. Metlagel Z, Kikkawa YS, Kikkawa M. Ruby-Helix: An implementation of helical image processing based on object-oriented scripting language. *J Struct Biol*, 2007, 157(1): 95-105
 51. Wang HW, Nogales E. An iterative Fourier-Bessel algorithm for reconstruction of helical structures with severe Bessel overlap. *J Struct Biol*, 2005, 149(1): 65-78
 52. Ramey VH, Wang HW, Nogales E. *Ab initio* reconstruction of helical samples with heterogeneity, disorder and coexisting symmetries. *J Struct Biol*, 2009, 167(2): 97-105
 53. Behrmann E, Tao G, Stokes DL, Egelman EH, Raunser S, Penczek PA. Real-space processing of helical filaments in SPARX. *J Struct Biol*, 2012, 177(2): 302-313
 54. Meng X, Zhao G, Zhang P. Structure of HIV-1 capsid assemblies by cryo-electron microscopy and iterative helical real-space reconstruction. *J Vis Exp*, 2011. DOI: 10.3791/3041
 55. Zhang Y, Wang W, Chen J, Zhang K, Gao F, Gao B, Zhang S, Dong M, Besenbacher F, Gong W, Zhang M, Sun F, Feng W. Structural insights into the intrinsic self-assembly of Par-3 N-terminal domain. *Structure*, 2013, 21(6): 997-1006
 56. Frank J, Radermacher M, Penczek P, Zhu J, Li Y, Ladjadj M, Leith A. SPIDER and WEB: Processing and visualization of images in 3D electron microscopy and related fields. *J Struct Biol*, 1996, 116(1): 190-199
 57. Chan KY, Gumbart J, McGreevy R, Watermeyer JM, Sewell BT, Schulten K. Symmetry-restrained flexible fitting for symmetric EM maps. *Structure*, 2011, 19(9): 1211-1218
 58. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K. Scalable molecular dynamics with NAMD. *J Comput Chem*, 2005, 26(16): 1781-1802

Electron Microscopy Reconstruction of Helical Assemblies

ZHANG Yan, SUN Fei

National Laboratory of Biomacromolecular, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China

This work was supported by grants from the National Natural Science Foundation of China (31100617) and the "973" Program of China (2011CB910301, 2014CB910700)

Received: Nov 29, 2013 **Accepted:** Dec 5, 2013

Corresponding author: SUN Fei, Tel: +86(10)64888582, Fax: +86(10)64888376, E-mail: feisun@ibp.ac.cn

Abstract: The cryo-electron microscopy reconstruction technique has been developed rapidly in recent years and becoming an important approach for structural studies of macromolecular complexes, and it includes three different methods, cryo-electron tomography, electron crystallography, and single particle analysis (SPA). Among them, the SPA technique is now becoming a powerful tool to determine near-atomic resolution structures of macromolecular complexes. However, for those biological macromolecular assemblies with helical symmetry, such as tobacco mosaic virus (TMV), microtubules, microfilament, human immunodeficiency virus 1 capsid protein and etc., it is not easy to obtain their three dimensional structures by using SPA because there are lots of math to deal with. In this paper, the authors described the helical reconstruction technique in details, including the basic mathematical description of helical assemblies, helical diffraction and its indexing, the relationship between Fourier transform of helical points array and that of real helical assemblies, and how to determine the helical parameters. Based on the above introduction, the authors reviewed two main helical reconstruction algorithms, the Fourier-Bessel reconstruction and the iterative helical real space reconstruction (IHRSR). At the end, they selected the helical assembly of Par-3 NTD (Par3 is a kind of protein related to cell polarity control and Par-3 NTD is its N-terminal domain) as an example to show a detailed helical reconstruction protocol using IHRSR algorithm.

Key Words: Helical assemblies; Cryo-electron microscopy; Three dimensional reconstruction; Fourier-Bessel transform; Iterative helical real space reconstruction

DOI: 10.3724/SP.J.1260.2013.30159