中 国 科 学
1987 年 3 月　　B 辑　　第 3 期

天花粉蛋白分子的三维结构*

潘克桢　林玉娟
傅珠玑　周康靖
蔡兆平　陈志伟　张永茂
（中国科学院福建物质结构研究所，福州）
董贻诚　吴　伸
马星奇　王耀萍
陈世芝　王家槐　张学军
（中国科学院生物物理研究所，北京）
倪朝周　张志明
夏宗芗　范肇昌　田庚元
（中国科学院上海有机化学研究所）

摘　　要

本文报道了我们最近建立的单斜晶胞内一个不对称单元中两个天花粉蛋白分子的模型，指出蛋白分子结构可以划分为大小两个结构域以及二级结构的一些特点，描述了独立单元中以一种非对称方式相关联的两个蛋白分子的差异及其相互作用，并解释了重原子衍生物的结合位置．

前文报道了天花粉蛋白分子的分子边界以及肽链走向[1,2]．我们又用四圆衍射仪把母体晶体和 Pt，Hg 二个重原子衍生物晶体的衍射数据扩展到 3.0 Å 分辨率；用迴摆照相法收集了母体晶体和 Pt，U，Hg 三个衍生物晶体的 2.6 Å 分辨率衍射数据．并分别计算母体相角和电子密度函数，建立了不对称单元中两个蛋白分子的结构模型．本文将着重报道天花粉蛋白的晶体和分子结构．

一、实　　验

1. 数据收集和处理

(1) 衍射仪法　3.0 Å 分辨率的衍射数据是在 PW1100 四圆衍射仪（荷兰 Philips）上收集的．母体晶体和 $K_2Pt(NO_2)_4$ 衍生物晶体的衍射数据从原有 3.5 Å 扩展到 3.0 Å，K_2HgI_4 衍

本文 1986 年 4 月 19 日收到．

* 单位名称按英文字母顺序．

下列同志参加了本工作：吴兹武、郑安、叶晓明、陈元柱、徐本杰（中国科学院福建物质结构研究所）、郝治平、马益林、孙孝先、唐天宝（中国科学院上海有机化学研究所）．

生物的衍射数据从 4.0Å 扩展到 3.0Å，其它实验条件同前文[1,2]. 表 1 列出母体晶体衍射数据不同分辨率的 R_{sym}.

表 1 不同分辨率下偏离因子 R_{sym}

分辨率范围	∞—3.5Å	3.5Å—3.0Å	∞—3.0Å
独立衍射点数	6331	3218	9549
R_{sym}	0.054	0.093	0.072

(2) 迴摆照相法 我们在国内首次应用迴摆照相机 (荷兰 Enraf-Nonius) 收集了一套 2.6Å 分辨率的衍射数据，所用 X 射线源是旋转阳极 X 光机(日本理学电机)，收集数据时采用平板照相盒. 母体晶体和 $K_2Pt(NO_2)_4$, UO_2Ac_2, K_2HgI_4 三个衍生物晶体的衍射数据均在同样实验条件下收集的. 表 2 列出数据收集的参数.

表 2 数据收集选择的参数

项 目	参 数
晶体大小	0.6 mm×0.4 mm×0.1 mm
射 线	CuKα (Ni 滤波), 150mA, 40kV
射线准直器孔径	0.6 mm
晶体到照相盒的距离	75 mm
晶体取向	i) b^* 安装法: 晶体的 b^* 轴垂直于入射线，且和仪器旋转轴平行，当 X 射线沿 c 轴方向入射时，取 $\phi = 0°$ ii) c^* 安装法: 晶体的 c^* 轴垂直于入射线，且和仪器旋转轴平行，当 X 射线沿 a 轴方向入射时，取 $\phi' = 0°$
晶体转动速率	3,000 s/°
每组照片收集范围	2°
每组底片数	3 张
底片来源	医疗用 X 光底片 II 型(天津感光胶片厂)
底片显定影温度	恒温 25℃

晶体取b^*安装法照相时，晶轴转动角度ϕ的范围为 0°—40°, 60°—100° 和 120°—160°，这样的收集范围已经包括反常散射在内的半个反射球的衍射数据，即可减少 60° 的收集范围，节省时间；又减少了蛋白晶体的辐射损伤，而全部数据损失只占 6% 左右(包括衍射的盲区在内). 为了弥补这一数据损失，又采用了 c^* 安装法收集数据，其 $\phi' = 0°$—32°.

衍射照片用 AD-1 微光密度计(美国 Nicolet) 测量强度，由于天津产 X 光底片对 X 射线的衰减率较差，在每一组照相中，放置 3 张底片于照相盒内，测量衍射强度只测第 1 和第 3 张，中间的一张作为射线衰减用，经过这样的处理，已能在第 1 和第 3 张底片(强度衰减比约为 2.25:1) 中测量到全部衍射点强度 (包括强衍射点). 数据处理用 MOSCO, 3D-SCALE 等程序在 PDP11/70 计算机上进行. 母体晶体的等效衍射点的平均强度的 R_{sym} 列于表 3.

母体晶体共有 14,264 个独立衍射点($I \geqslant 36(I)$),其平均 $F_P = 19.20$.

表 3　不同分辨率的偏离因子 R_{sym}(照相法)

分辨率范围(Å)	衍射点数	R_{sym}
100.000—10.000	205	0.031
10.000—7.071	519	0.037
7.071—5.774	711	0.039
5.774—5.000	808	0.040
5.000—4.472	942	0.037
4.472—4.083	1035	0.042
4.083—3.780	1072	0.048
3.780—3.535	1146	0.048
3.535—3.333	1197	0.053
3.333—3.162	1232	0.068
3.162—3.015	1320	0.066
3.015—2.887	1248	0.084
2.887—2.774	1299	0.088
2.774—2.673	1197	0.101
2.673—2.582	1280	0.112

2. 电子密度函数的计算

我们用两种方法收集的 3.0 Å 和 2.6 Å 分辨率衍射数据分别计算了电子密度函数.

3.0 Å 分辨率的母体相角计算是分壳层进行的. 参加计算的 Pt, Hg 二个重原子衍生物的参数是以 4 Å 分辨率的参数作为初始值,然后用 4—3.0 Å 壳层的数据进行修正. 修正后的 R (F_{HLE})[1] 为 0.498 (K_2HgI_4) 和 0.599 ($K_2Pt(NO_2)_4$). 用修正后的重原子参数计算 4—3.0 Å 壳层的联合概率相角,此套相角的平均品质因子为 0.669. 用此套相角和原 4 Å 三个衍生物 (Pt, U, Hg) 计算所得的相角合并,得到一套完整的 3.0 Å 分辨率的母体相角,并以此计算了 3.0 Å 分辨率的电子密度函数.

2.6 Å 分辨率的母体相角计算是用 Pt, U, Hg 三个重原子衍生物数据计算的. 重原子参数以 4 Å 分辨率的参数作为初始值,用全部 2.6 Å 衍射数据进行修正. 修正后的 R_{crys}[1)]分别为 0.638(Pt), 0.701(U) 和 0.556(Hg). 用修正后的重原子参数计算 2.6 Å 分辨率的联合概率相角,其平均品质因子为 0.583, 并以此计算 2.6 Å 分辨率的电子密度函数.

3. 分子模型的建立

将所得的 3.0 Å 和 2.6 Å 分辨率的两套电子密度图,按 1 Å = 2 cm 比例画在透明薄膜上,用 Richards[3] 光学叠合装置建立不对称单元中两个独立分子(分别以分子 A, 分子 B 表示)的模型. 此外,还以所建立的分子 A 的结构模型,应用前文[2]得到的线性变换关系,并经适当调整后得到分子 B 的坐标.

二、结果和讨论

1. 电子密度图和分子模型

在开始建立分子模型时,虽然一级结构工作尚未完成,但已有大部分一级结构信息可供利

1) $R_{crys} = \dfrac{\Sigma|F_{PHC} - F_{PH}|}{\Sigma|F_{PH} - F_P|}$

用．我们运用结构化学知识将已知部分的氨基酸顺序以及已经得到的肽链走向和电子密度图配合。一级结构全部完成时[4,5]，我们又参照全部氨基酸顺序，核对分子模型，发现它能够基本上和一级结构吻合．[3]图 1 为不对称单元中两个天花粉蛋白分子的模型．

用两种数据收集方法得到的 3.0 Å 和 2.6Å 分辨率数据计算的电子密度图，其结果是大体一致的．蛋白分子的主链与电子密度图符合较好，大部分侧链也有相应的密度．

在天花粉蛋白分子的氨基酸顺序中，有二处出现三个连续芳香环残基，即 Tyr88—Phe89—Phe90 和 Phe147—Tyr 148—Tyr 149，在 3.0 Å 或 2.6 Å 分辨率的电子密度图上都能见到这一特征的电子密度分布．蛋白分子其余大部分芳香环残基，都有相应的侧链密度．9 个 Phe 残基与电子密度图符合较好；13 个 Tyr 残基中，除 114，118 位较差，81 位最差外，其余均与图符合较好．1 个 His 残基的侧链密度符合较好．但是，1 个 Trp 残基，在 3.0 Å 图中分子 A 可以找到对应的电子密度，分子 B 不理想；而在 2.6 Å 图中则符合较差．此外，9 个 Pro 残基均符合较好．12 个 Gly 残基除 127 位密度稍大外，其余均无大的侧链密度．

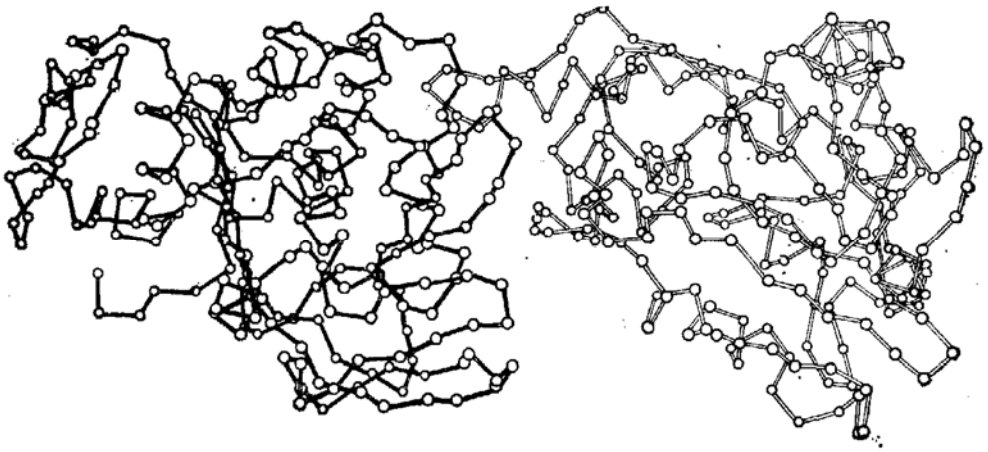
3.0 Å 分辨率的电子密度图比较清晰，蛋白分子与溶剂界线明显，主链密度较清楚，大部分侧链可以找到对应的密度．但是，在自由肽段主链断裂、表现模糊，侧链密度也不太好．特别在 Pt_1，Hg_1 和 Hg_2 位置附近有很重的电子密度无法解释，这和相角误差有关．

2.6 Å 图的主链密度比较清楚，大部分的侧链可以找到对应的密度，但是有些地方蛋白分子和溶剂的界线不太清楚，需经仔细辨认才能区分．有几处主链密度断裂(特别是在分子表面的自由肽段)，此外也有一些无法解释的密度，总的来讲背景较高，这也是由于相角误差所致．

统观 4 Å，3 Å 和 2.6 Å 分辨率的电子密度图，随着分辨率的升高，图的质量明显下降，这可能是由于有的重原子衍生物偏离因子较大和同晶度不好所致．然而，由于我们在建立分子模型时，依据了 3.0 Å 和 2.6 Å 分辨率的电子密度图，并且还参考了 4 Å 分辨率的肽链走向，这样就收到了互相印证、互相补充的效果．

2. 蛋白分子的描述

图 2 表示不对称单元中两个蛋白分子(仅含 C^α 原子)的主链结构(每个分子由 234 个氨基酸残基组成)．



分子 A 分子 B

图 2 不对称单元中两个蛋白分子的主链结构图

由图可见，天花粉蛋白分子结构可以看作由两个结构域组成，大结构域从N端开始，约含169个氨基酸残基；小结构域包括了C端在内的约65个残基．两个结构域间除有一个主链连接点外，还有一个由邻近C端的β_{13}链和大结构域的β_2，β_3链形成的β折叠结构，这是两个结构域间重要的非共价键的相互作用．整个蛋白分子的结构有如儿童玩具不倒翁，小结构域是不倒翁的头部，大结构域为不倒翁的胖躯体，而C端（包括β_{13}链）的主链有如一条围巾绕着不倒翁的脖子．

表4给出天花粉蛋白分子的二级结构单元．

表4　天花粉蛋白分子的二级结构单元

序　　号	α　螺　旋	β　链
1	Ser11—Leu25	Val2—Leu6
2	Lys104—Leu112	Arg29—Tyr32
3	Tyr114—Tyr118	Asp33—Ile37
4	Gln122—Ile132	Ile49—Tyr55
5	Ala138—Phe147	Ala56—Asp62
6	Tyr149—Leu158	Ala74—Ile78
7	Arg170—Phe186	Met79—Ala83
8	Asn213—Ser219	Glu97—Ala99
9		Glu119—Leu121
10		Thr165—Glu167
11		Ser198—Leu203
12		Ser204—Ile209
13		Gln227—Asn228
合　计	84 a. a.	61 a.a.
占全部a.a.的百分比	36%	26%

天花粉蛋白分子所含的8段α螺旋相对集中在蛋白分子的中央，其中特别是螺旋α_5，几乎整个地内埋于分子中心，由13条β链所构成的4个β折叠层则分布在蛋白分子的表面，这是天花粉蛋白分子结构的显著特点．

天花粉蛋白的二级结构中有两处特别引人注目．一是螺旋α_7在Gly179处有一个明显的弯折，其弯折达70°左右，这么大的弯折角在蛋白结构中是很罕见的，以致可以把这个中间断裂三对氢键（Gln176，Gln177，Ile178的羰基氧和Ser180，Arg181，Val182的酰胺氮）的α螺旋看作是由两段α螺旋（α_{7A}和α_{7B}）所组成，处于中间的Gly179是公共残基，在两段的螺旋中都形成氢键．另一处是由β_1，β_4，…，β_8六条链所组成的典型左手扭曲的β折叠层，这是一个混合型β折叠层．正如通常混合型β折叠层那样，由于平行的和反平行的折叠层结构参数的不同，往往折叠层中间主体是一种类型，两边的两条链可以是另一种类型．在天花粉蛋白分子中，中间的β_4，…，β_7是反平行的四条链，其中，β_4和β_5，β_6和β_7构成典型的反平行发夹结构；边上的β_1和β_8则分别相对于β_4和β_7作平行取向．有趣的是，尽管它符合一般的蛋白质结构规律，但是β_6链的最后一个残基Ile78的酰胺氮却未能和β_5链的Glu58的羰基氧形成氢键，从而使β_5和β_6链之间被撕开了一个口子，最终使整个大β折叠层显示出比通常要大的左

手扭曲.

天花粉蛋白分子结构还有一个使人感兴趣的特点即由三条伸展的肽链形成的三个环状肽段聚集在蛋白分子的一侧，它位于大β折叠层附近. 环1是由β_3到β_4之间约11个残基(Arg38—Ala48)构成，环2是由β_5—β_6间约11个残基(Val63—Asn73)组成，而环3则由β_7—β_8间约13个残基(Gly84—Thr96)构成. 其中环3以典型的右手方式联结两条平行的肽链β_7与β_8，环2则好像以右手方式扭转起来的双螺旋结构. 环与环之间，环与其他肽段之间，存在着盐键、氢键和疏水作用，因而这三个环能够相互聚集在一起，并使得三个这样大的环状肽段仍具有相当确定的构象.

天花粉蛋白分子大约一半是疏水残基，它们大部分内埋在分子中. 在分子内部，并非所有疏水残基都聚集在一起，但是比较集中的有两处，大小结构域各有一个稍大的疏水区. 在大结构域中主要是由螺旋α_1和β_1，β_4，β_5链的内侧以及螺旋α_5羧端面向大β折叠层的一侧围聚而成，它包含Phe4，Phe17等芳香环和其他疏水残基. 在小结构域中，则是在螺旋α_9和β_{11}，β_{12}折叠层之间的界面上，以Trp200为中心而聚集起来的疏水区.

从分子模型可以看出，两个蛋白分子的表面都存在一个正电荷相对集中区，这个区域位于大小结构域交界处的表面凹槽附近，这里聚集了一些Arg，Lys的残基，这些带正电荷的残基，有的可能与附近带负电荷的Asp，Glu残基形成盐键，但是，大部分还是游离的. 此外，还存在一些盐键，有的在分子内部，有的在分子表面，有的则在分子间形成.

蛋白分子的氨基酸顺序中有一个连续的Gly—Gly顺序，位于螺旋α_4的中间，并未造成螺旋结构参数的异常，这是有趣的. 分子内9个Phe残基，全部内埋或指向分子的内部.

3. 不对称单元中两个蛋白分子的差异及其相互作用

从总体看，两个分子的立体结构基本上是一样的. 我们从目前的电子密度图上也看到一些两个分子的构象差别，比较明显的是肽链羧端若干残基和连接二级结构的一些无规环状肽段的构象差，个别螺旋的长度也许略有不同. 我们注意到己糖激酶，它比天花粉蛋白分子大整整一倍，两个不对称分子之间的构象差主要限于其相互接触的界面上，即使这些区域，两个分子的叠合位移也只在1Å左右[6]. 因此，在现有的精度下，过分精细的比较显然是无意义的.

如图3所示，不对称单元中两个天花粉蛋白分子间相互作用力主要集中在沿晶胞c轴方向的两个分子的交界处. 它们有二个交界. 一个交界主要是分子A的大片β折叠层(即β_1，β_4，$\cdots$，β_8)和分子B的β_2，β_3，β_{13}所组成的β折叠层(包括C末端)，β_3—β_4的环状肽段之间的相互作用. 在这一交界处，有许多残基在可以形成盐键、氢键或疏水键的距离内相互交错在一起. 另一个交界是在分子A沿c轴的相反方向，主要是分子A的β_{11}，β_{12}折叠层和分子B的螺旋α_4—α_5间的变形"拱桥"构型肽段的疏水残基之间的相互作用. 此外，分子A的α_7—β_{11}，β_{12}—α_8间的伸展肽段和分子B的螺旋α_4氮端、α_6—β_{10}间的伸展肽段的几个亲水残基相互交错，形成分子间的氢键.

沿晶胞a轴方向的分子间相互作用是由晶胞的2次轴对称转换相关的两个蛋白分子之间的作用，如图3所示，形成了一些氢键和少数疏水残基间的相互作用.

沿b轴方向，如图3，是蛋白分子A或B自身(经2_1螺旋轴对称转换)之间的相互作用，可能形成极少数的氢键和个别疏水残基接触的疏水作用，例如分子A的螺旋α_3，β_9，β_{10}反平行β折叠层和分子A′(经对称转换而来)的β_2链、螺旋α_{7B}羧端之间的相互作用.

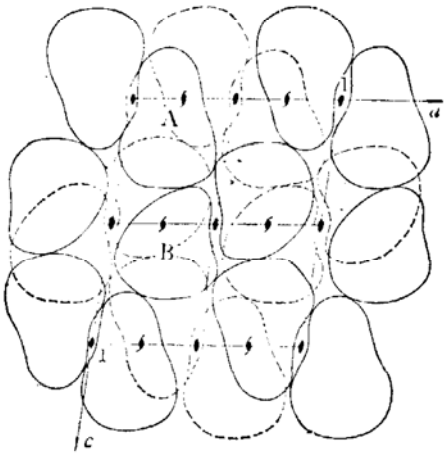


图 3　天花粉蛋白分子在晶胞中排列的投影图
（沿 b 轴方向投影）

由于这些分子间的相互作用，使得不对称单元中两个天花粉蛋白分子所处的晶体学环境十分不同，因而两个蛋白分子有较为奇特的相互取向．在天花粉蛋白晶体培养的实验中，晶体的外形沿 c 轴方向总比其他方向长得多．这与两个蛋白分子沿 c 轴方向相互作用相当强有密切关系．

4. 重原子结合位置

这个分子模型还很好地解释求解相角时所使用的重原子衍生物的结合位置．$Pt(NO_2)_4^{2-}$ 有三套结合位置：Pt_1 除与分子 B 的 Met79 结合外，分子 A 的 N 端的正电荷也加强了 $Pt(NO_2)_4^{2-}$ 这个带负电荷的络离子的结合．Pt_2 只与分子 A 的 Met79 结合．由于这两二个 Pt 原子的结合环境不同，因而其相对占有率亦不同，Pt_1 结合最强，其相对占有率为 1.09，而 Pt_2 只有 0.51．Pt_3 结合在分子 B 的 C 端的 Met233，由于距 2 次对称轴较近，因此，其相对占有率最低，仅为 0.29．分子 A 的 C 端距 2 次对称轴更近，所以没有 Pt 的结合．

HgI_4^{2-} 亦有三套结合位置．Hg_1 和 Hg_2 分别被两个蛋白分子中处于小结构域表面凹陷区的 Lys172 及 Lys205 所环抱，因此，这二个重原子的相对占有率都很高，分别为 0.91 和 1.00．而 Hg3 则位于两个蛋白分子之间，仅与分子 B 的 Lys30 结合，其相对占有率较低，为 0.60．

UO_2^{2+} 有二套结合位置，U_1 和分子 A 的 Glu196 以及经 2 次轴对称转换的分子 A′ 的 Ser39 相结合；而 U_2 仅与分子 B 的 Tyr88 结合，因此，其相对占有率相差很大，分别为 1.01 和 0.27．

英国牛津大学 Stuart 和 Todd 在 2.6 Å 分辨率数据收集过程中提出宝贵意见并赠送迴摆照相的数据处理程序以及其他有关程序，特此致谢．

参　考　文　献

[1] 潘克桢等，中国科学 B 辑，1982，6：521．
[2] 潘克桢等，中国科学 B 辑，1985，8：718—723．

[3] Richards, F. M., *J. Mol. Biol.*, **37**(1968), 225.
[4] Wang Yu, et al., *International Symposium on Organic Chemistry of Medicinal Natural Products (IUPAC)*, Nov. 1985, Shanghai, China, Abstr. A08.
[5] 顾梓伟等，化学学报，**42** (1984)，943—945.
[6] Steitz, T. A., et al., *J. Mol. Biol.*, **104**(1976), 197.