

用 MXAN 拟合钩端螺旋体去甲酰化酶 XANES 谱*

郭晓云^{1,2} 储旺盛² 龚为民³ 董宇辉² 谢亚宁² 杨飞飞²

Maurizio Benfatto⁴ 吴自玉^{2;1)}

1 (中国科学技术大学生命科学学院 合肥 230026)

2 (中国科学院高能物理研究所 北京 100049)

3 (中国科学院生物物理所 北京 100101)

4 (Istituto Nazionale di Fisica Nucleare, Laboratori Nazionali di Frascati, P.O. Box 13, 00044, Frascati, Italy)

摘要 钩端螺旋体去甲酰化酶 (*Leptospira interrogans* PDF) 是一种重要的含锌金属蛋白酶, 对钩端螺旋体这一广泛存在的致病菌的蛋白合成起着关键的催化作用, 是一个很好的药物设计靶蛋白. 本文测试了 LiPDF 在 pH3.0 的溶液状态下的 X 射线吸收近边结构 (XANES: X-Ray Absorption Near-Edge Structure) 谱, 利用以从头计算 (ab. initio) 的多重散射 (Multiple Scattering) 为基础的 MXAN 方法确定金属蛋白活性中心的精细结构. 研究发现结合合适的初始结构模型, 可以更好地重现 LiPDF 蛋白的 XANES 曲线, 从而能够得到更加准确的结构参数. 活性中心的精细结构为理解 LiPDF 的 pH 依赖的催化活性提供了结构基础.

关键词 钩端螺旋体去甲酰化酶 金属蛋白 XANES MXAN 程序包

1 引言

蛋白质的生物合成均起始于甲硫氨酸. 在大多数情况下, N 端的甲硫氨酸在成熟蛋白中被去除. 一般认为只有在细菌中, N 端甲硫氨酸在蛋白合成之前被甲酰化. 在翻译后加工中, N 端甲硫氨酸只有先在去甲酰化酶 (peptide deformylase, PDF) 作用下去甲酰化, 才可以被正确切除. 研究表明 PDF 对于细菌存活是必需的. 而在真核生物中, 目前发现蛋白的翻译后加工工程不涉及 PDF 的作用. 因此, PDF 是一个很好的药物设计靶蛋白. 钩端螺旋体去甲酰化酶 (*Leptospira interrogans* PDF) 是一种重要的含锌金属蛋白酶, 对钩端螺旋体这一广泛存在的致病菌的蛋白合成起着关键的催化作用. LiPDF 酶催化活性受到反应体系中多种因素, 尤其是 pH 值的影响.

金属蛋白大约占已知蛋白总数的四分之一到三分之一, 利用金属的氧化还原性质和配位化学特性来实现其特定的生物学功能, 这个过程往往伴随着金

属离子周围微弱的结构变化. X 射线吸收谱学 (XAS: X-Ray Absorption Spectroscopy) 是一种具有亚原子分辨率的局域结构研究方法, 具有不受样品形态限制、对金属离子具有选择性等突出优点, 是研究金属蛋白中金属活性中心精细结构的理想技术^[1-6]. XAS 曲线可以分为两个部分, 一般将吸收边以上 50eV 以内的吸收曲线称为 X 射线吸收近边结构 (X-ray Absorption Near-Edge Structure, XANES), 而吸收边以上 50eV 一直延伸到 1keV 的区域称为扩展 X 射线吸收精细结构谱 (Extended X-ray Absorption Fine Structure, EXAFS). 对于低浓度的蛋白样品, EXAFS 谱的信噪比通常很差, 使得其应用受到了很大的限制. 而 XANES 谱不仅可以给出丰富的电子结构信息, 还能够提供金属格点周围立体的几何结构信息. 因此, 人们一直尝试克服理论计算 XANES 谱的复杂性, 直接从 XANES 提取出吸收原子周围的立体结构.

MXAN 程序是近几年来才发展起来的在从头计算 (ab initio) 的多重散射 (Multiple Scattering) 理论

2006 - 05 - 17 收稿

* 杰出青年基金(10125523)和国家自然科学基金重大交叉研究项目(10490191, 10490193)资助

1) E-mail: wuzy@mail.ihep.ac.cn

框架下利用X射线吸收近边结构重建吸收原子周围局域结构的定量拟合程序^[6-11]. 势函数的计算采用Muffin-Tin近似^[7, 12], Muffin-Tin半径根据Norman规则选取^[13], 即最近邻原子间的Muffin-Tin半径可根据原子序数为 Z 的原子包含所有的 Z 个电子的半径比率来确定. 交换关联势根据激发光电子自能的局域密度近似来确定, 使用的是一个复数形式的光学势^[14], 包括光电子经历的净的Coulomb势 V_{coul} (不依赖于能量的改变又被称为“静态”势函数)和自能 $\Sigma(E)$ (类似于基态计算中的交换相互作用势, 其具体的函数形式依赖于所采取的近似). 此光学势也依赖于 N 粒子的终态效应, 即依赖于激发原子的内层空穴态和原子/分子构型. MXAN既可以用和能量有关的 X_{α} 势, 也可以用Hedin-Lundqvist(H-L)和能量无关的势^[15], 一般采用H-L复势 (complex potential) 或者H-L势的实部 (real potential). 对于分子体系的拟合, H-L势的复数部分会导致主通道跃迁振幅的过度损失, 而主通道包含所有的结构信息^[16, 17]. 因此MXAN使用实的势函数部分计算出的吸收截面, 再利用具有 $\Gamma = \Gamma_c + \Gamma(E)$ 形式的Lorentzian加宽函数进行卷积. MXAN程序是由意大利Frascati国家实验室的科学家开发的, 并在不断改进中, 包括选用新的内坐标系进行计算等. 本文采用了H-L势的实部, 对LiPDF的pH3.0的溶液样品的XANES谱进行了拟合, 拟合得到的计算谱与实验谱吻合得很好. 比较拟合前后金属活性中心的精细结构, 不仅为理解LiPDF的酸性环境下的低催化活性提供了结构基础, 还为研究药物设计提供了结构信息.

2 实验

LiPDF分子量约20.2kD, 蛋白溶液样品的制备在中国科学院高能物理研究所多学科中心生物大分子实验站的生化实验室进行. LiPDF的大肠杆菌中高效表达菌株由中国科技大学生命科学学院提供, 携带质粒pET22b-LiPDF大肠杆菌BL21(DE3)菌株按照^[18]表达纯化得到蛋白纯品. 再将蛋白纯品用浓度为50mmol/L的pH3.0的Glycine-HCl缓冲液稀释 10^3 倍后再将蛋白浓缩到60mg/mL.

LiPDF的锌 K 边XANES数据是在常温下在北京同步辐射实验室(BSRF)1W1B光束线上采集的. 储存环运行能量为2.2GeV, 流强范围为80—120mA. Wiggler引出的同步辐射光经硅双晶单色器分离出单色光, 双晶单色器失谐30%以抑制高次谐波. 数据采集采用荧光模式, 采用Cu滤波片, 前电离室缓慢通

入25%氩气/75%氮气的混合气体, 荧光电离室采用100%的氩气.

溶液样品测试时封装如下: 在厚度为1.5mm的铝板上开一个直径为5mm的孔, 完全清洗干净后将溶液样品采用很薄的塑料薄膜封在其中, 同时保证各处溶液厚度的均匀并没有气泡. 在正式采谱前后, 通过快速扫描模式收集样品的XANES谱, 并进行比较以确认X射线中对样品的损伤情况.

3 结果与讨论

比较采谱前后通过快速扫描收集的XANES谱可知两条曲线吻合得相当好, 而且吸收边的位置也没有移动, 这说明在扫描时间内X光的照射对样品损伤较小, 没有导致锌周围局域结构的变化. 实际上, 为了确认X光长时间照射对样品的损伤, 在实验前我们专门对此蛋白的溶液样品进行了多次扫描, 并将每次得到的谱线叠加在一起进行比较, 没有发现可以归因于光致蛋白变性的谱图变化, 而且从蛋白溶液样品的外部形态上也没有发现变化, 因此认为实验得到的谱图是可以代表样品中锌离子周围原始的局域结构的.

为了能够利用MXAN程序重现LiPDF的锌 K 边XANES谱, 需要确定合适的初始结构模型. 选择pH3.0时和甲酸结合状态下的复合物的结构(PDB代码1Y6H). 此蛋白在pH3.0的晶体结构中是个二聚体, A链做初始模型试拟合过程中, 主吸收峰的双峰结构很难出现, 而B链结构作为初始模型拟合时则很容易得到双峰结构. 另一方面, 1Y6H是pH3.0时和甲酸结合状态下的复合物的结构, 原生态中和锌离子配位的催化水分子在此结构中被甲酸提供的两个氧原子所取代, 锌离子也从原生态的四配位转换为复合物中的五配位, 如图1所示. 为了更准确的表示原生态结构,

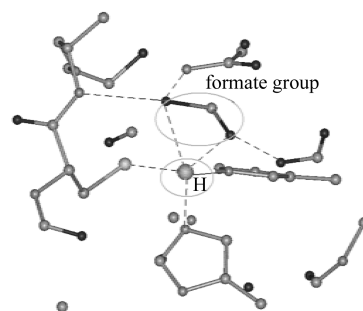


图 1 用球棍模型显示的1Y6H的B链的金属活性中心局域结构

重要基团之间的成键用虚线显示; 图中椭圆突出的甲酸根的两个氧原子取代了原生态中和锌离子配位的催化水分子, 用圆圈突出的小球代表的是五配位的锌离子.

参考大肠杆菌 PDF 的晶体学结构 (PDB 代码 1BS5), 将与锌配位的两个氧原子用两个水分子代替, 分别距离 Zn 原子 2.31Å 和 3.90Å^[19].

使用 MXAN 程序对吸收谱近边结构的拟合首先应该确定合适的原子团簇 (cluster) 尺寸. 因为 X 射线激发出来的光电子具有一定的电子平均自由程, 对于 X 射线吸收谱, 这个从芯态激发出来的光电子有效的散射半径为光电子平均自由程的一半. 因此, 半径过小的原子团簇很难用来重现 XANES 谱, 但如果原子团簇包含的原子数目过多, 又会极大地增加计算量. 这里可以不断增加原子团簇中的原子数, 计算出相应的 XANES 谱, 直到 XANES 谱不再出现变化为止, 以确定原子团簇的大小, 如图 2 所示. 由图中可以看出, 15 个原子的原子团簇就已经给出了几乎所有的主要结构. 这个原子团簇包含的最主要配位体就是两个组氨酸的咪唑环、半胱氨酸最近邻的几个原子和 3 个水分子 (其中最近的水分子与 Zn 配位并且参与催化反应), 从而可以说明近邻原子的散射给出了吸收谱的最主要贡献. 最后 3 个原子团簇的理论谱已经非常一致, 表示本文所用的含 50 个原子左右的原子团簇 (距离金属离子半径 5.9Å) 已经能够很好的用来重现 LiPDF 的 XANES 谱.

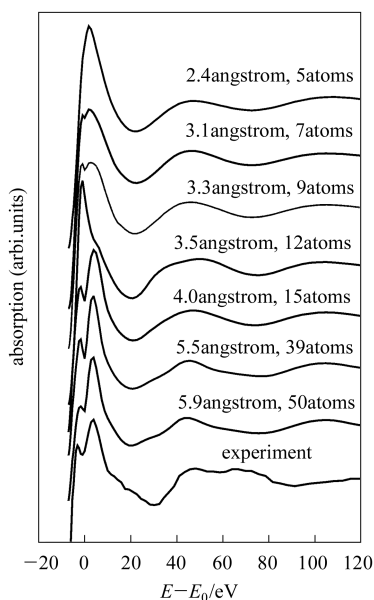


图 2 理论计算的 LiPDF 蛋白的锌 K 边 XANES 谱随原子团簇尺寸的变化曲线

MXAN 程序拟合可以分为两个过程: 1) 电子结构参数优化; 2) 几何结构参数优化. 前一个过程优化的电子参数主要有 Fermi 能级、Muffin-Tin 近似中相邻原子势函数重叠的百分比 (Overlap factor)、Muffin-Tin 零点的位置、用于描述空穴寿命和实验分辨率对

XANES 谱有展宽影响的 Lorentzian 函数的宽度常数 (broadening) Γ 以及光电子非弹性散射过程对 XANES 谱也有展宽影响的并与能量相关的 Lorentzian 函数宽度常数 $\Gamma(E)$ 的位置 (Epl1) 和强度 (Apl1). 几何结构参数主要优化的是近邻配位原子和吸收原子之间的键长和键角, 从给定的初始结构模型出发, 不断修正初始结构模型, 使得理论计算谱不断趋近于实验谱. 在实际拟合过程中, 不断交替使用上述两个优化过程, 使得理论计算谱和实验谱之间的均方差 (S^2) 达到最小.

图 3 给出了利用 MXAN 程序最终得到的计算谱和实验谱的比较, 可以看出实验谱和计算谱吻合得很好. 相对应的 LiPDF 蛋白活性中心锌格点周围的局域精细结构与初始给出的由结构晶体学数据得到的局域结构的比较如图 4 所示.

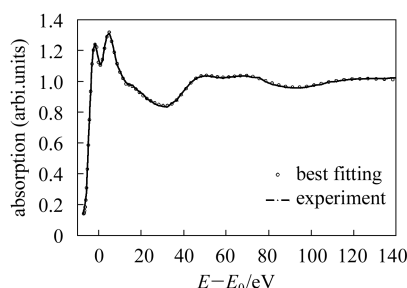


图 3 LiPDF 的 pH3.0 的溶液样品的锌 K 边 XANES 实验谱和拟合谱的比较

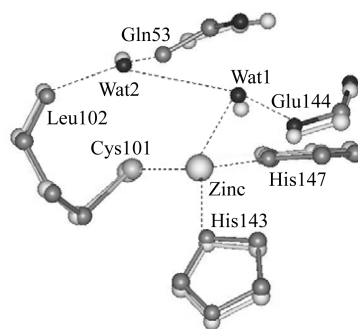


图 4 LiPDF 的初始结构和通过 MXAN 拟合得到的最终局域结构的叠合
重要基团之间的成键用虚线显示; 图中用浅色的球棍模型显示并置于图像底层的是初始结构, 拟合得到的最终结构用深色球棍模型显示并位于图像顶层.

将拟合前后活性中心周围的局域结构叠合在一起, 可以发现其精细结构有了变化, 特别是催化水分子 Wat1 和催化亚基 Glu144. 表 1 对比了金属离子的 4 个配位键长 (分别与 His147 和 His143 主链上的 N 原子、Cys101 侧链上的 S 原子和催化水分子 Wat1) 以及催化水分子与 Glu144 之间的氢键键长在拟合前后的数值.

表 1 初始模型给出的键长和通过拟合 XANES 谱最终得到的键长

| 成键原子对 | 初始模型/Å | 拟合结果/Å |
|-------------|--------|------------|
| Zn_His147 | 2.15 | 2.04±0.001 |
| Zn_His143 | 2.17 | 2.08±0.001 |
| Zn_Cys101 | 2.4 | 2.24±0.004 |
| Zn_Wat1 | 2.31 | 2.63±0.006 |
| Wat1_Glu144 | 2.67 | 2.51±0.019 |

从表1和图4可以看出, 锌离子与His147和His143五元环上的氮原子、一个催化水分子Wat1, Cys101侧链上的硫原子这4个原子形成四配位, 空间上呈4面体构型. 催化水分子Wat1还与催化亚基Glu144侧链上的氧原子形成一个氢键, 催化过程中最关键的步骤就是Wat1在Glu144帮助下对底物氮末端乙酰基的羰基碳的亲核攻击, 脱去一个质子给底物碳原子. 与初始模型相比, 锌的四配位距离有3个都缩短了0.09Å—0.16Å, 而Zn_Wat1的距离却增加高达0.32Å. 由于此初始模型是来自于pH3.0下的复合物的晶体学数据, 因此可以抛弃pH对结构的影响. 认为其配位键长的较大变化可能来源于3方面, 一是状态不同, 分别是晶体和溶液状态, 但这并非主要原因, 因为晶体状态下蛋白的含水率大于40%; 二是配位数目的影响, 初始模型是在复合物五配位基础上修改的, 金属与His147, His143和Cys101的配位键长更是直接来自五配位时的数据, 因此认为配位数目的增加会影响金属离子原先配位距离; 更重要的一个原因是通过X射线晶体学得到的数据误差较大(1Y6H分辨率2.2Å, 误差达0.2Å), 很难准确地描述蛋白活性中心的局域精细结构, 尤其是活动性比较大的水分子.

与LiPDF干粉样品的MXAN拟合相比^[19](采用EcPDF 1BS5的晶体结构为初始模型), 新的初始结构模型使理论谱和实验谱吻合得更好. 干粉结构锌离子的4个配位分别为2.08Å, 2.08Å, 2.21Å和2.24Å, His147, His143和Cys101的拟合结果差别不大, 而

Wat1与锌离子的配位距离则相差非常大, 达到0.39Å. 一方面可能是由于初始模型的改进, 另一方面来自其本身物性: 干粉样品重新溶解在中性环境催化活性仍然很高, 而pH3.0的LiPDF则几乎没有催化活性^[20]. 因此可以猜测正是由于这个水分子与锌离子的配位距离会随着环境pH变化而改变, 导致了催化活性随pH变化. 当溶液从中性变为酸性时, 外界环境中的负电荷减少, 对催化水分子Wat1的质子的静电吸引减弱, 表现在水分子与锌离子的配位距离增长, 导致了Wat1不能脱去质子并将这个质子转移到底物羰基碳上, 从而丧失催化活性.

4 结论

蛋白质晶体学给出的数据很难准确地描述金属活动中心周围的精细结构, 而X射线吸收谱对金属原子周围的局域结构及其敏感, 可以作为蛋白质晶体学方法的补充, 在金属蛋白质结构和功能的研究中发挥重要的作用. 本文利用XANES方法, 从蛋白质晶体学给出的结构模型出发, 结合多重散射理论求解得到了LiPDF蛋白在pH3.0环境下的金属活性中心周围的局域精细结构. 研究发现使用合适的初始结构模型, 可以更好地重现LiPDF蛋白的XANES曲线, 从而能够得到更加准确的结构参数. 本文还比较了干粉样品和pH3.0溶液样品金属活动中心的局域结构, 并讨论了酸性环境下催化活性丧失的原因. 这个结果不仅有助于对这一类pH依赖的催化活性的理解, 还为研究药物设计提供了结构信息. 更深一步地理解催化活性对pH值的依赖关系还需要研究更多pH环境下的金属蛋白结构, 这项工作正在进行并将在以后的文章中讨论.

对谢亚宁老师、胡天斗老师、刘涛老师以及马思玄同学给予的帮助表示诚挚的谢意.

参考文献(References)

- Hasnain S S, Hodgson K O. *J. Synchrotron Rad.*, 1999, **6**: 852
- Koningsberger D C, Prins R. *X-Ray Absorption: Principles, Applications, Techniques of EXAFS, SEXAFS and XANES*. New York: John Wiley & Sons Inc., 1988. 257—320
- Scott R A. *Structural and Resonance Techniques in Biological Research*. Rousseau D L, ed., Orlando, Florida: Academic Press, 1984. 295
- Powers L. *Biochim. Biophys. Acta*, 1982, **1**: 683
- Rehr J J, Ankudinov A L. *J. Synchrotron Rad.*, 2003, **10**: 43
- Natoli C R, Benfatto M, Della Longa S et al. *J. Synchrotron Rad.*, 2003, **10**: 26
- Natoli C R, Benfatto M. A Unifying Scheme of Interpretation of X-Ray Absorption Spectra Based on the Multiple Scattering Theory. In: Lagarde P, Raoux D, Petiau J. *EXAFS and Near Edge Structure IV*, Volume 1. Paris: Les Editions de Physique, 1986. 11—23
- Della Longa S, Arcovito A, Benfatto M et al. *Physical Review Letters*, 2001, **87**: 155501(1—4)
- Benfatto M, Congiu-Castellano A, Della Longa S et al. *J. Synchrotron Rad.*, 2001, **8**: 267—269

- 10 Benfatto M, Della Longa S. *J. Synchrotron Rad.*, 2001, **8**: 1087—1094
- 11 Benfatto M, Della Long S, Natoli C R. *J. Synchrotron Rad.*, 2003, **10**: 51—57
- 12 Durham P, Pendry J B, Hodges C H. *Comput. Phys. Commun.*, 1982, **25**: 193—200
- 13 Norman J G. *Mol. Phys.*, 1976, **31**: 1191—1198
- 14 Tyson T A, Hodgson K O, Benfatto M. *Phys. Rev.*, 1992, **B46**: 5997—6019
- 15 Hedin L, Lundqvist S. *Solid State Phys.*, 1969, **23**: 1—181
- 16 Benfatto M, Solera J A, Garcia J. *Phys. Rev.*, 1997, **B56**: 2447—2452
- 17 Rehr J J, Albers R C. *Rev. Mod. Phys.*, 2000, **72**: 621—654
- 18 LI Y, REN S, GONG W. *Acta Cryst. D.*, 2002, **58**: 846—848
- 19 LI Shu-Jun, ZHOU Zhao-Cai, WU Zi-Yu et al. *J. Synchrotron Rad.*, 2005, **12**: 111—114
- 20 LI Y, CHEN Z, GONG W. *Biochem Biophys Res Commun.*, 2002, **295**(4): 884—889

MXAN Analysis of the XANES Energy Region of LiPDF^{*}

GUO Xiao-Yun^{1,2} CHU Wang-Sheng² GONG Wei-Min³ DONG Yu-Hui² XIE Ya-Ning²
YANG Fei-Fei² Maurizio Benfatto⁴ WU Zi-Yu^{2;1)}

1 (School of Life Science, University of Science and Technology of China, Hefei 230026, China)

2 (Institute of High Energy Physics, CAS, Beijing 100049, China)

3 (National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China)

4 (Istituto Nazionale di Fisica Nucleare, Laboratori Nazionali di Frascati, P.O. Box 13, 00044, Frascati, Italy)

Abstract *Leptospira interrogans* peptide deformylase (LiPDF) is an important Zinc metalloprotein and plays a key role in protein synthesis of *Leptospira interrogans*, an ubiquitous environmental bacterium that can cause the infections of animals and humans. It is an attractive target for the design of new antibiotics since the deformyling apparently does not occur in the cytoplasm of eucaryotic cells. In this paper, we probed the protein in pH3.0 solution by X-ray absorption near-edge structure (XANES) spectroscopy. We quantitatively determined the local structure around the metal center using MXAN procedure, which is based on Multiple Scattering theory. The experimental XANES spectrum of LiPDF has been reconstructed perfectly using suitable initial structure mode, and the corresponding local structure of the active center provides structure information for understanding the pH-dependent enzymatic activity.

Key words LiPDF, metalloprotein, XANES, MXAN package

Received 17 May 2006

^{*} Supported by Outstanding Youth Fund (10125523) and Key Important Project of the National Natural Science Foundation of China (10490191, 10490193)

1) E-mail: wuzy@mail.ihep.ac.cn