

# 分子动力学研究亚铁血红素激活 蛋白转录激活机理

赵 熹, 黄旭日, 孙家钟

(吉林大学理论化学研究所, 理论化学计算国家重点实验室, 长春 130023)

**摘要** 以 3 种亚铁血红素激活蛋白 (Heme activator protein, HAP)-DNA 复合物 (野生型 HAP1-wt, Ser63/Arg63 突变 HAP1-18 和 Ser63/Gly63 突变 HAP1-PC7) 为对象对亚铁血红素激活蛋白的转录激活机理进行了分子动力学研究. 对 3 个复合物分子动力学轨迹的比较性分析显示, 涉及到上游活化序列 (Upstream activation sequences, UAS) 识别的蛋白质-DNA 相互作用分布与实验观测到的 3 种蛋白转录活性一致. 进一步对 3 个复合物进行柔性分析显示, 3 个 DNA 分子具有相似的柔性, 而又有所不同, 特别在涉及 UAS 识别的 N-端和 Zn2Cys6 结构域前部有明显的柔性差异. 蛋白质柔性的差别导致不同的蛋白质-DNA 相互作用. 因此亚铁血红素激活蛋白的 N-端和 Zn2Cys6 结构域前部的柔性大小能够调节亚铁血红素激活蛋白转录激活功能.

**关键词** 分子动力学; 转录激活; 柔性; 亚铁血红素激活蛋白

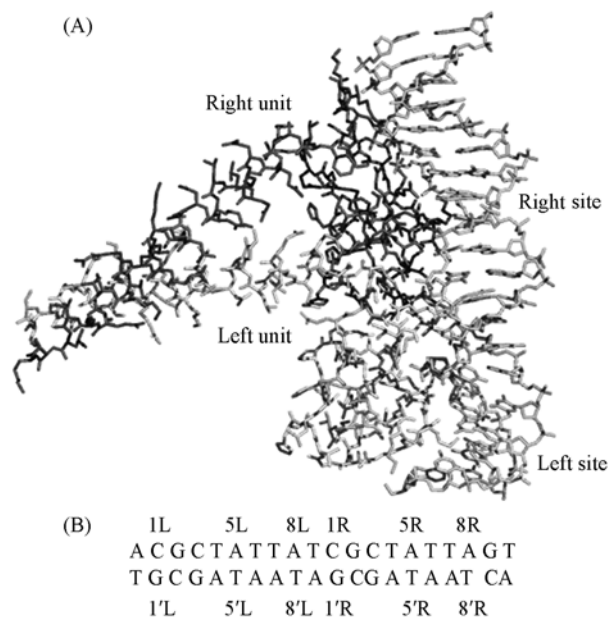
**中图分类号** O641

**文献标识码** A

**文章编号** 0251-0790(2008)04-0819-05

亚铁血红素激活蛋白 (Heme activator protein, HAP) 是一个含有进行 DNA 识别 N-端 Zn2Cys6 结构域和一个 C-端转录活化域的转录因子<sup>[1~3]</sup>. 这种转录因子主要调节涉及磷酸化以及修补过程中的基因<sup>[4~9]</sup>, 调节是通过结合这些基因的上游活化序列 (Upstream activation sequences, UAS) 实现的. 亚铁血红素激活蛋白以同型二聚体 (Homodimer) 的形式识别同一 DNA 分子上的两段上游活化序列; 其 N-端 Zn2Cys6 结构域识别上游活化序列; 而 C-端的卷曲螺旋及连接部分结构 (Coiled-coil and linker) 对连接两段 UAS 之间的碱基对进行识别 (图 1).

亚铁血红素激活蛋白的活化机制一直不清楚. 文献<sup>[10~12]</sup>运用突变和 X 射线衍射结晶方法发现亚铁血红素激活蛋白的 63 号残基对该蛋白的活化性质至关重要; 与初始的铁血红素激活蛋白相比, 两个突变蛋白虽然保持相似的 DNA 结合能力, 但 (Ser63→Gly63) 突变的亚铁血红素激活蛋白失去转录活性, 而 (Ser63→Arg63) 突变的亚铁血红素激活蛋白则明显提高了活性. 虽然它们通过对静态的 X 射线衍射结构分析指出这种活性的变化是蛋白与 DNA 之间相互作用变化的结果, 但具体的机理仍然不清楚的. 图 1 (A)



**Fig. 1 Overview of the HAP1-wt-UASCYC7 complex**

(A) Schematic diagram of the HAP1-wt-UASCYC7 complex. The right and left subunits are labeled, respectively, and the corresponding DNA half-sites also are labeled. (B) The base pair HAPI binding sites from UASCYC7 in crystal structure and MD simulations. The numbering scheme used in the text is indicated above or below the corresponding base. The UAS of both targets are indicated in bold.

收稿日期: 2007-01-31.

基金项目: 国家自然科学基金 (批准号: 20103003)、教育部骨干教师基金和吉林省杰出青年基金资助.

联系人简介: 黄旭日, 男, 博士, 教授, 博士生导师, 主要从事理论化学研究. E-mail: huangxr@mail.jlu.edu.cn

为亚铁血红素激活蛋白-DNA 的分子图形。

本文选择野生型和两个突变型亚铁血红素激活蛋白-DNA 复合物(Ser63/Gly63 HAP1-pc7 和 Ser63/Arg63 HAP1-18)的 X 射线衍射晶体结构作为研究对象,应用分子动力学方法发现亚铁血红素激活蛋白 63 号残基突变能导致其识别 UAS 的 N-端 Zn2Cys6 结构域柔性的明显变化,而三个复合物中 DNA 则呈现相似的柔性.因此,N-端 Zn2Cys6 结构域柔性对于调节亚铁血红素激活蛋白活化性质至关重要.图 1(B)为 DNA 碱基对示意图.

## 1 计算方法

### 1.1 分子动力学模拟

3 种亚铁血红素激活蛋白-DNA 复合物的 X 射线衍射晶体结构(野生型: PDB 代号 1HWT; Ser63→Gly63 突变: PDB 代号 1QP9; Ser63→Arg63 突变: PDB 代号 2HAP)作为分子动力学(Molecular dynamics)模拟的初始结构.在常温常压下(300 K, 1 个标准大气压),对这三个体系完成了 3 次独立充分含水的分子动力学模拟,模拟时间均为 6.0 ns,共计 18.0 ns.所用力场是 AMBER98<sup>[13,14]</sup>.软件包是 GROMACS3.21<sup>[15,16]</sup>,采用 SPC 水模型<sup>[17]</sup>.

用 LINCS<sup>[18]</sup>算法约束所有键长,体系原子的初速度由 Maxwellian 分布给出.系统温度通过由溶剂和溶质分别与参考温度的热浴耦合来保持<sup>[19]</sup>,温度耦合常数为 0.1 ps.系统压强通过与外部压强浴耦合实现,压强耦合常数为 0.5 ps.范德华和库仑相互作用的截断半径均为 1.4 nm,短程邻近截断距离(rlist)为 0.8 nm.

系统模拟过程:首先建立具有平衡 SPC 水分子的矩形的盒子,然后将 3 个亚铁血红素激活蛋白-DNA 复合物分别插入 3 个盒中,体系与盒子边的最小距离为 1.0 nm.为保持整个体系的总电荷为零,盒子中加入适当的 Na<sup>+</sup>离子,除去相应位置的水分子.准备好的体系都使用了周期化条件.用最速下降法对体系进行能量最小化,直到最大作用力小于 100 kJ/(mol·nm)为止,然后运行 200 ps 的平衡动力学.之后开始常温常压下的自由分子动力学模拟,步长为 2.0 fs.

此外,为了保证 Zn<sup>2+</sup>与 Cystin(胱氨酸)正确的配位状态,通过施加距离约束算法维持 Zn<sup>2+</sup>-CYS 和 Zn<sup>2+</sup>-Zn<sup>2+</sup>配位键的正常键长,力常数为 1000 kJ/(mol·nm<sup>2</sup>).

### 1.2 熵的计算

熵是衡量分子柔性大小的有效参数,在本文中,我们采用 Schafer 等<sup>[20,21]</sup>提出的熵变计算方法,即熵变与分子动力学模拟确定的质量协方差矩阵有关:

$$\Delta S \approx \frac{k_B}{2} \ln \det \left[ 1 + \frac{k_B T e^2}{\hbar^2} M_\sigma \right] \quad (1)$$

式中  $M_\sigma$  代表质量协方差矩阵本征值.

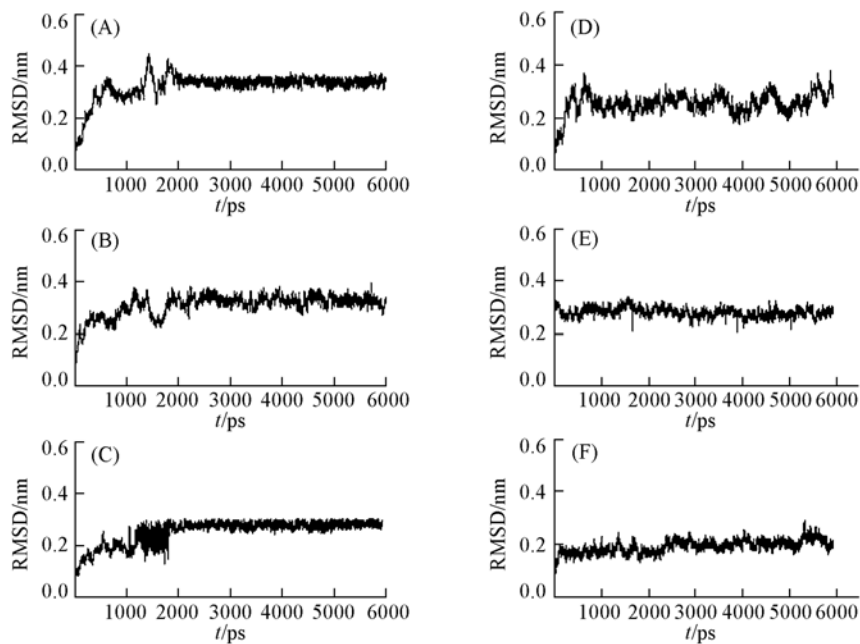
## 2 结果与讨论

### 2.1 动力学体系的平衡

3 个体系中的蛋白质和 DNA 骨架原子( $\alpha$  碳原子和磷原子)的均方根偏差(Root-mean-square deviations, RMSD)随模拟时间变化的趋势分别见图 2[(A)~(B)]和[(D)~(F)].图 2 清楚地显示了这 3 个复合物中的亚铁血红素激活蛋白质与 DNA 结构在 2.0 ns 左右达到动态平衡<sup>[22]</sup>.3 个蛋白(野生型 HAP1-wt, Ser63→Arg63 HAP-1-18 和 Ser63→Gly63 HAP-1-PC7)的平均 RMSD 分别为 0.31, 0.29 和 0.33 nm.

### 2.2 蛋白质-DNA 的相互作用

通过对比文献[10~12]的结果,3 种亚铁血红素激活蛋白与 DNA 上的两个 UAS 之间的氢键在系统平衡(2.0 ns~6.0 ns)基础之上被分析.分子模拟结果与 X 射线衍射结构测量的氢键受体与给体的距离列于表 1(标准氢键受体与给体的距离为 0.35 nm).由表 1 实验数据可见,虽然亚铁血红素激活蛋白质以同型二聚体形式结合含有两个 UAS(CGC)的 DNA,但它只能有效地识别其中之一.分子模拟



**Fig. 2 Time evolution of RMSD of MD structure with respect to the corresponding minimized structures**

The protein backbone heavy-atoms of the HAPI-wt (A), HAPI-18 (B) and HAP-PC7 (C), respectively. All the heavy-atoms of the three bound DNA: HAPI-wt (D), HAPI-18 (E) and HAPI-PC7 (F), respectively.

(表 1) 也得到这个结果, 即 3 种亚铁血红素激活蛋白二聚体都只能识别一个 UAS 片断上的 3 个碱基对 (即有效识别从 C1R...G1'R 到 C3R...G3'R 碱基对). 因此可推断亚铁血红素激活蛋白二聚体只能有一个蛋白起到转录活化功能 (即处于左侧的蛋白质分子). 表 1 也显示, 对于野生型 HAP-wt 和突变 (Ser63/Arg63) HAPI-18 都能借助氢键识别 UAS 的三个碱基对 (C1L...G1'L 到 C3L...G3'L), 而另一种突变亚铁血红素激活蛋白 (Ser63/Gly63) HAPI-PC7 仅能识别 UAS 的前 2 个碱基对 (即 C1L...G1'L 到 G2 L...C2'L), 因此它没有转录活性. 突变 (Ser63/Arg63) HAPI-18 可以与 UAS 形成 8 个氢键, 而野生型 HAP-1-18 只形成 4 个氢键, 因此突变 (Ser63/Arg63) HAPI-18 比野生型有更高的转录活性. 值得注

**Table 1 Hydrogen-bonding interactions between proteins and DNA molecules in three protein-DNA complexes\***

Protein	Hydrogen bond	HAPI-18	HAPI-wt	HAPI-PC7
Right protein	Arg70RO...C1RN4	0.285 (0.297 ± 0.024)	0.291 (0.315 ± 0.036)	0.279 (0.306 ± 0.031)
	Lys71RO...C2'RN4	0.278 (0.749 ± 0.052)	0.318 (0.276 ± 0.012)	0.273 (0.278 ± 0.012)
	Lys71RNz...G2RN7	0.266 (0.291 ± 0.016)	0.295 (0.336 ± 0.099)	0.281 (0.316 ± 0.016)
	Lys73RNz...C2'RO2P	0.260 (0.546 ± 0.082)	0.283 (0.470 ± 0.085)	0.268 (0.427 ± 0.082)
	Cys74RN...C3'RO1P	0.270 (0.431 ± 0.053)	0.450 (0.538 ± 0.038)	0.280 (0.386 ± 0.049)
	Lys76RNz...G3'RO1P	0.707 (0.719 ± 0.053)	0.370 (2.981 ± 0.099)	0.266 (0.417 ± 0.062)
Left protein	Arg63LNh2...G3'LN7	0.307 (0.502 ± 0.032)	—	—
	Arg63LNh1...G3'LO1P	0.310 (0.275 ± 0.017)	—	—
	Arg70LO...C1LN4	0.283 (0.301 ± 0.042)	0.344 (0.285 ± 0.019)	0.281 (0.290 ± 0.053)
	Lys71LO...G3'LN7	0.300 (0.341 ± 0.034)	0.433 (0.371 ± 0.025)	0.393 (0.410 ± 0.039)
	Ly71LO...C2'LN4	0.289 (0.283 ± 0.015)	0.329 (0.297 ± 0.032)	0.271 (0.380 ± 0.023)
	Lys71LNz...G2LN7	0.267 (0.297 ± 0.023)	0.294 (0.383 ± 0.061)	0.282 (0.290 ± 0.027)
	Lys71LNz...C3LN4	0.287 (0.372 ± 0.054)	0.347 (0.587 ± 0.045)	0.287 (0.753 ± 0.076)
	Lys71LNz...G2LO6	0.281 (0.470 ± 0.057)	0.301 (0.430 ± 0.057)	0.342 (0.387 ± 0.058)
	Lys71LO...G1'LO6	0.472 (0.345 ± 0.034)	0.458 (0.271 ± 0.018)	0.431 (0.511 ± 0.032)
	Lys73LNz...C2'LO2p	0.266 (0.396 ± 0.090)	0.350 (0.520 ± 0.032)	0.454 (1.223 ± 0.014)
	Lys73LNz...C3'LO2p	0.266 (0.539 ± 0.013)	0.388 (0.930 ± 0.092)	0.479 (0.575 ± 0.059)
	Cys74LO...G3'LO1P	0.342 (0.333 ± 0.019)	0.393 (0.545 ± 0.056)	0.485 (0.556 ± 0.035)
	Cys74LN...G3'LO1 (2) P	0.279 (0.310 ± 0.018)	0.306 (0.322 ± 0.043)	0.474 (0.504 ± 0.019)
	Lys76LNz...G3'LO2P	0.493 (0.611 ± 0.057)	0.390 (0.679 ± 0.022)	0.634 (0.446 ± 0.067)

\* Average non-bonding distance (nm) and their standard deviations (both in parenthesis) for the contacts sites of HAP-DNA complexes. The values of X-ray structure of HAP-DNA (PDB code: 2HAP, 1HWT and 1QP9).

意的是:对这3种亚铁血红素激活蛋白除了63号残基,其余参加UAS识别的残基都相同( ARG70, LYS71, LYS73 和 CYS74),但是它们对UAS识别能力去不一致,这是它们具有不同柔性的体现.

### 2.3 柔性变化与转录激活机理

通过计算3个复合物体系中的蛋白与DNA骨架原子均方根浮动(RMSF; root mean square fluctuations)揭示了3个体系的柔性特点. 3个体系中左侧蛋白的 $\alpha$ 碳原子( $C_\alpha$ )和DNA的磷原子(P)的RMSF值分别见图3(A)和(B)中. 沿着整个亚铁血红素激活蛋白二级结构链,除了(Ser63/Gly63)突变HAP1-PC7的Zn2Cys6结构域的一个小片断(残基为82~87)以外,(Ser63/Arg63)HAP1-18突变和野生型亚铁血红素激活蛋白的骨架柔性要高于HAP1-PC7. 特别是对于涉及UAS识别的N-端和Zn2Cys6结构域前部(残基为55~76),HAP1-18有最高的柔性,而HAP1-PC7则具有最低的柔性;最高不超过0.12 nm. 另一方面,沿着碱基顺序,3个体系中DNA分子呈现相似的骨架柔性,RMSF平均值不超过0.20 nm. 特别是两个UAS片断,三个体系柔性非常相似.

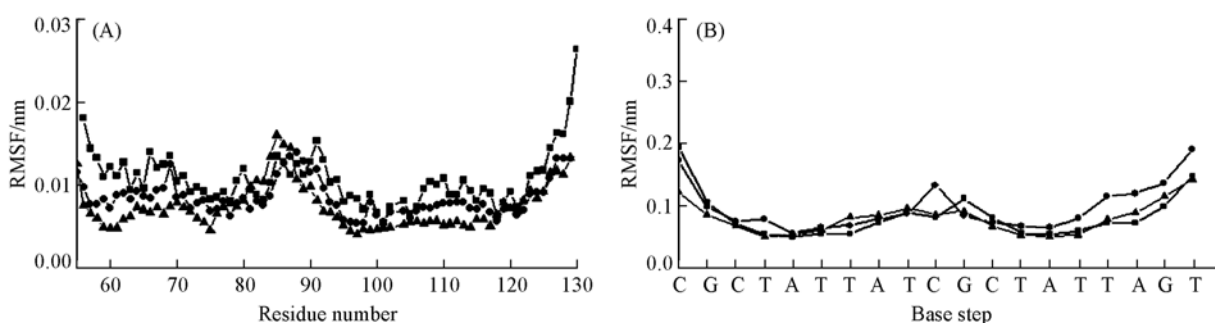


Fig. 3 RMSF of the  $C_\alpha$  atoms around the average MD structures(A), and P atoms in single chain of DNA(ACGCTATTATCGCTATTAGT) (B)

(A) ● HAP1-wt; ■ HAP1-18; ▲ HAP1-PC7. (B) ● HAP1-wt; ■ HAP1-18; ▲ HAP1-PC7.

通过分子内部熵(Intramolecular entropy)可以对分子体系的柔性大小进行计算. 计算结果(表2)表明,3种亚铁血红素激活蛋白的活性顺序为(Ser63/Arg63)HAP1-18 > HAP1-wt > (Ser63→Gly63)HAP1-PC7,而3个对应的DNA由于熵值相似,因此有相似的柔性. 这个计算结果与RMSF分析一致. 因此可以推断,对63号残基的突变可以改变亚铁血红素激活蛋白N-端和Zn2Cys6结构域前部(残基:55~76)的柔性,但这种突变对结合的DNA分子构象影响较小. 因此亚铁血红素激活蛋白与DNA的相互作用主要依赖于亚铁血红素激活蛋白的柔性变化.

Table 2 Intramolecular entropies (kJ/mol) for the backbone of three HAPs and the corresponding bound DNA

Species	HAP1-18	HAP1-wt	HAP1-PC7
Protein Backbone	24102.59	23760.75	22820.96
DNA backbone	7792.23	7793.36	7820.03

结合上述对柔性所进行的分析,3个体系的蛋白质-DNA相互作用差异主要是由3种亚铁血红素激活蛋白的柔性差异导致的,相比于(Ser63→Gly63)HAP1-PC7,(Ser63→Arg63)HAP1-18和野生型亚铁血红素激活蛋白在N-端和Zn2Cys6结构域前部具有高的柔性,能够充分地UAS的3个碱基对接触,因此能够与UAS建立有效的分子识别. 总之,亚铁血红素激活蛋白的N-端和Zn2Cys6结构域前部的柔性大小对调节亚铁血红素激活蛋白的转录活性至关重要.

### 参 考 文 献

- [1] Marmorstein R., Carey M., Ptashne M., *et al.*. Nature[J], 1992, **356**: 408—414
- [2] Schjerling P., Holmberg S.. Nucleic Acids Res. [J], 1996, **24**: 4599—4606
- [3] Johnston M.. Microbiol. Rev. [J], 1987, **51**: 458—476
- [4] Lalonde B., Arcangioli B., Guarente L.. Mol. Cell Biol. [J], 1986, **6**: 4690—4696
- [5] Pfeife K., Arcangioli B., Guarente L.. Cell [J], 1987, **49**: 9—18

- [ 6 ] Pfeifek K. , Prezant T. , Guarente L. . Cell[J] , 1987 , **49** : 19—27
- [ 7 ] Zitomer R. S. , Sellers J. W. , Mecarter D. W. , *et al.* . Mol. Cell Biol. [J] , 1987 , **7** : 2212—2220
- [ 8 ] Winkler H. , Adam G. , Mattes E. , *et al.* . EMBO. J. [J] , 1988 , **7** : 1799—1804
- [ 9 ] Lodi T. , Guiard B. . Mol. Cell Biol. [J] , 1991 , **11** : 3762—3772
- [ 10 ] King D. A. , Zhang L. , Guatente L. , *et al.* . Nature Structural Biology[J] , 1999 , **6** : 22—27
- [ 11 ] King D. A. , Zhang Li. , Guatente L. , *et al.* . Nature Structural Biology[J] , 1999 , **1** : 64—71
- [ 12 ] Amanda K. , Lukens D. A. , King D. A. , *et al.* . Nucleic Acids Res. [J] , 2000 , **28** : 3853—3863
- [ 13 ] Cornell W. D. , Cieplak P. , Bayly C. I. , *et al.* . J. Am. Chem. Soc. [J] , 1995 , **117** : 5179—5197
- [ 14 ] Cheatham T. E. , Cieplak P. , Kollman P. A. . J. Biomol. Struct. Dynam. [J] , 1999 , **16** : 845—862
- [ 15 ] Berendsen H. J. C. , Vanderspoel D. , Van D. R. . Comp. Phys. Commun. [J] , 1995 , **91** : 43—56
- [ 16 ] Lindahl E. , Hess B. , Vanderspoel D. . J. Mol. Mod. [J] , 2001 , **7** : 306—317
- [ 17 ] Berendsen H. J. C. , Postma J. P. M. , Van Cunstern W. F. , *et al.* . Pullman B. Ed. Intermolecular Forces[M] , Dordrecht: Reidel Publishing Company, 1981: 331—342
- [ 18 ] Hess B. , Bekker H. , Berendsen H. J. C. , *et al.* . J. Comp. Chem. [J] , 1997 , **18** : 1463—1472
- [ 19 ] Berendsen H. J. C. , Postma J. , Van G. W. , *et al.* . J. Chem. Phys. [J] , 1984 , **81** : 3684—3690
- [ 20 ] Schafer H. , Mark A. E. , Van G. W. . J. Chem. Phys. [J] , 2000 , **113** (18) : 7809—7817
- [ 21 ] Andricioaei I. , Karplus M. . J. Chem. Phys. [J] , 2001 , **115**(14) : 6289—6292
- [ 22 ] WANG Song(王嵩) , HUANG Xu-Ri(黄旭日) , GAO Xue-Feng(高雪峰) , *et al.* . Chem. J. Chinese Universities(高等学校化学学报)[J] , 2006 , **27**(3) : 535—537

## Molecular Dynamics Study of Transcriptional Activation Mechanism of Heme Activator Protein

ZHAO Xi , HUANG Xu-Ri\* , SUN Chia-Chung

(State Key Laboratory of Theoretical and Computational Chemistry, Institute of Theoretical Chemistry,  
Jilin University, Changchun 130023, China)

**Abstract** A molecular dynamics(MD) simulations for three kinds of protein(heme activator protein: HAP)-DNA complexes crystal structures were performed: the HAP1-wt, two HAP1 mutants (HAP1-PC7: Ser63/Gly63; HAP1-18: Ser63/Arg63) with the aim of investigating the mechanism of HAP transcriptional activation. Comparative analyses of MD structures for the three HAP-DNA complexes reveal that the key protein-DNA interactions involving the recognition of UAS *i. e* CGC are different in three complexes as the experimental observations. Further analyses reveal that three HAPs exhibit different flexibilities, relative to very similar conformations of three bound DNA. It is found that the difference of flexibilities in three HAPs results in diversities in conformations of N-term Arm and Zn2Cys6 Binuclear Cluster involving DNA recognition, causing varieties of protein-DNA interactions. According to these results, the flexibility of N-term and Zn2Cys6 Binuclear Cluster in HAP can play a crucial role in regulating transcriptional activation, which can directly lead to the alternative protein-DNA interactions.

**Keywords** Molecular dynamics; Transcriptional activation; Flexibility; Heme activator protein

(Ed. : Y, I)