

全 α 蛋白质体系能量转变的分子动力学模拟*

姜舟婷^{1**} 孙婷婷² 王亚楠¹ 陈昱东¹ 何志兴¹ 徐 鹏¹

(¹ 中国计量学院理学院 杭州 310018) (² 浙江工商大学信息与电子工程学院 杭州 310018)

摘 要 采用简化的 NHP 模型并考虑蛋白质链的二级结构信息,对全 α 型蛋白质链进行构建.利用分子动力学模拟研究全 α 蛋白质体系势能的变化情况,得出键伸缩和弯曲能量随蛋白质的链长线性增加;单键的伸缩能和弯曲能随着温度的升高而线性增加,且伸缩能比弯曲能随温度的增加更为明显的结论.键扭转能随二级结构 α 螺旋数目的增加而线性降低,随温度的上升,扭转能变化呈现先增加后减少的趋势.非键能的大小随疏水残基在蛋白质组分中的比例增加而降低,但线性关系不明显.键扭转能和非键能都存在转变温度.蛋白质链的结构是由多种势能协同作用的结果,键能和非键能的相互竞争以及组成蛋白质链的氨基酸亲疏水性和二级结构组分都会影响蛋白质体系的能量,这些结果对深入理解蛋白质的结构提供理论依据.

关键词 全 α 蛋白质, 分子动力学模拟, 能量转变

蛋白质是一个复杂的软凝聚态体系,是生命体的重要组成部分.研究表明,蛋白质存在相同的折叠类型(即二级结构),主要包括 α 螺旋(α -helix), β 折叠(β -sheet)和 β 转角(β -turn),基于此折叠识别,可进一步构建其三级结构^[1].由于蛋白质三维空间结构测定手段的局限性,目前越来越多的学者致力于从蛋白质二级结构出发预测蛋白质的三级结构,这些研究不仅为蛋白质三级结构的预测提供了有效的途径,也为解释蛋白质结构和生物功能之间的关系提供有力的手段^[2~4].在蛋白质结构分类数据库(SCOP)中,蛋白质类型主要归于全 α 、全 β 、 α/β 、 $\alpha + \beta$ 四大类^[5].本文研究的蛋白质在 SCOP 数据库中皆归类于全 α 类型蛋白质,其二级拓扑结构大部分由 α 螺旋构成.螺旋结构在生物大分子中普遍存在,特别在受限环境下,会发生螺旋结构的转变^[6],因此对全 α 蛋白质链的研究具有重要的意义.

蛋白质晶体 X 射线衍射和 NMR 测定技术是当今蛋白质空间结构的主要测定方法^[7,8],然而其耗资巨大、蛋白质分离纯化的技术要求高,且有些蛋白质晶体(如膜蛋白)难以培养,测定的周期较长等原因都制约着蛋白质工程的发展^[9,10].随着近代物理学、化学和生命科学的交叉以及计算机技术的进步,人们开始用理论和计算机模拟的

方法来研究蛋白质体系.在基于多维核磁共振或 X 射线晶体衍射实验数据的结构修正中,分子模拟已成为常规手段^[11~13].计算机模拟和理论分析的研究中,对蛋白质分子的建模从最简单的 HP 格点模型^[14],过渡到非格点的粗粒化模型^[15],进而发展到描述所有原子和物理相互作用的全原子模型并开发了相关软件.粗粒化模型虽不能给出蛋白质折叠过程中所有的原子细节,但能在更长的时间尺度执行计算,这是了解整个折叠过程的动力学行为和热力学特征并挖掘潜在物理机制的关键.利用计算机模拟各种势能场中的蛋白质结构和性质在一些领域获得与实验相同的结果^[14~18].

蛋白质结构研究的理论基础是 20 世纪 70 年代由 Anfinsen 提出的“自组装学说”,认为蛋白质的折叠结构在一定条件下是热力学最稳定的状态.提出这一假说的实验依据是许多蛋白质可以在体外自发的由完全去折叠态变为自然折叠态^[19].本文选择 50 条不同链长的全 α 型蛋白质,对组成蛋白质的 20 种氨基酸进行简化分类,同时在模型中考虑蛋白质的二级结构组成,对组成蛋白质的各部分能量随外界温度的转变进行分析.这些结果对蛋白质结构的预测提供理论依据,从而加深对生命体的了解.

* 2013-05-06 收稿,2013-07-10 修稿;国家自然科学基金(基金号 21204078,20904047,11202201)、浙江省自然科学基金(基金号 Y6110304,LY12B04003)和浙江省科技厅钱江人才项目(项目号 2010R10022)资助项目.

** 通讯联系人,E-mail:z. jiang@cjlu. edu. cn

doi:10. 3724/SP. J. 1105. 2014. 13169

1 模型和计算方法

蛋白质非格点粗粒化的结构如图 1 所示,组成蛋白质的氨基酸由 4 个原子团组成, NH, C α , CO 和表示侧链原子团的 R,不同氨基酸的亲疏水性在 R 中体现. 在我们以前的工作中已经尝试依据 Miyazawa-Jernigan 的相互作用矩阵^[20],把 20 种氨基酸残基分为中性残基 (neutral, N), 疏水残

基 (hydrophobic, H) 和亲水残基 (hydrophilic, P) 并得到很好的模拟结果^[21~24]. 在本文工作中,我们采用非格点的 NHP 模型,同时考虑氨基酸残基的二级结构属性,对蛋白质链进行构建,用简化的氨基酸残基组合来表示蛋白质的序列,从而简化蛋白质能量的计算. 本文选取 50 条不同链长的全 α 型蛋白质 (PDB code 见表 1), 蛋白质链的初始构型在高温下随机产生,并首先进行能量最小化.

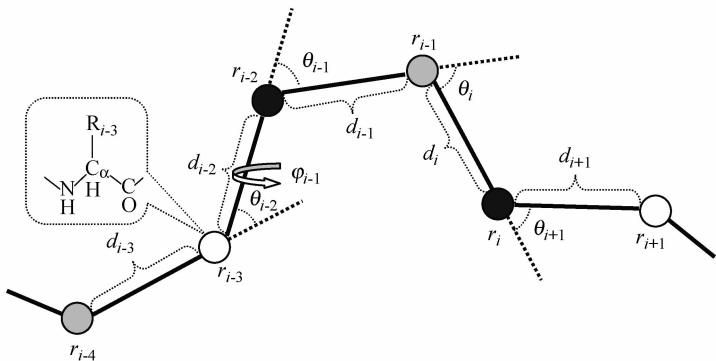


Fig. 1 Schematic representation of protein with NHP model
($\circ$, $\bullet$ and $\bullet$ represent H, N, P residues, respectively.)

Table 1 PDB codes of 50 all- α proteins

PDB code	Chain length	N	H	P	α helix	β sheet	others	PDB code	Chain length	N	H	P	α helix	β sheet	others
1ALA	316	77	106	133	220	0	96	1RRO	108	21	30	57	57	4	47
1BBL	37	12	10	15	23	0	14	1UTG	70	12	25	33	53	0	17
1BP2	123	18	44	61	67	10	46	1YCC	108	31	30	47	52	7	49
1CCR	111	30	30	51	53	0	58	1YEA	112	32	28	52	54	0	58
1EA8	140	26	46	68	127	0	13	2C2C	112	33	31	48	59	0	53
1ECA	136	41	45	50	116	0	20	2CDV	107	37	24	46	19	6	82
1ECD	136	41	45	50	116	0	20	2CTS	437	104	158	175	315	7	115
1ECO	136	41	45	50	116	0	20	2CY3	118	41	27	50	26	4	88
1F63	154	43	49	62	133	0	21	2CYP	293	62	93	138	155	16	122
1FCS	154	43	50	61	121	0	33	2END	137	30	49	58	61	0	76
1FHA	172	34	57	81	125	0	47	2FAL	146	43	45	58	112	0	34
1FIP	73	14	27	32	59	0	14	2HBG	147	56	44	47	120	0	27
1H96	167	37	55	75	135	0	32	2LHB	149	39	50	60	113	0	36
1HBG	147	56	44	47	120	0	27	2MHR	118	24	42	52	76	0	42
1IFA	158	21	66	71	91	0	67	2PDE	43	12	14	17	22	0	21
1LE4	139	25	45	69	126	0	13	2WRP	104	20	35	49	78	0	26
1LH1	153	41	53	59	126	0	27	3C2C	112	33	31	48	59	0	53
1LH2	153	41	53	59	126	0	27	451C	82	23	23	36	42	0	40
1LPE	144	29	46	69	128	0	16	4BP2	64	11	25	28	39	0	25
1MBA	146	43	45	58	112	0	34	4CPV	108	34	30	44	70	0	38
1MBC	153	45	48	60	121	0	32	4ICB	76	9	24	43	56	2	18
1MBD	153	45	48	60	121	0	32	4MBN	153	45	48	60	121	0	32
1MBS	153	44	46	63	121	0	32	5CPV	108	34	30	44	60	6	42
1PPA	121	23	44	54	66	4	51	5CYT	103	29	30	44	52	0	51
1RCB	129	31	41	57	80	9	40	5TNC	161	32	50	79	106	10	45

假设蛋白质链处在真空环境中,体系势能由键接相互作用(键伸缩能、弯曲能、扭转能)和非键相互作用构成,力场参数参见文献[25], $E_{\text{total}} = E_{\text{stretch}} + E_{\text{bend}} + E_{\text{torsion}} + E_{\text{LJ}}$,具体公式如下:

(1)键伸缩能(bond-stretching term)存在于每2个邻近的由共价键联系的氨基酸之间,由于化学键长在其平衡值附近呈小幅度震荡形成的.

$$E_{\text{stretch}} = \sum_{i=2}^N \frac{1}{2} k_b (r - \sigma)^2 \tag{1}$$

(2)键弯曲能(bond-bending term)存在于每3个邻近的氨基酸之间,由化学键的键角在其平衡值附近呈小幅震荡形成.

$$E_{\text{bend}} = \sum_{i=3}^N \frac{1}{2} k_\theta (\theta - \theta_0)^2 \tag{2}$$

(3)分子链中连续4个原子所形成的二面角(dihedral angel)较为松软,易于扭动,形成键扭转能(torsion angle term).

$$E_{\text{torsion}} = \sum_{i=4}^N \left[-A \cos^6 \left(\frac{\phi - \phi_\alpha}{2} \right) - B \cos^6 \left(\frac{\phi - \phi_\beta}{2} \right) \right] \tag{3}$$

(4)非键相互作用形式为 Lennard-Jones (LJ) 势能.

$$E_{\text{LJ}} = \sum_{i=1}^N \sum_{j=1}^N 4\varepsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \Lambda \left(\frac{\sigma}{r_{ij}} \right)^6 \right] \tag{4}$$

$(j - i \geq 4)$

上述公式中的相关参数: $k_b = 100\varepsilon_h/\sigma^2, k_\theta = 13.33\varepsilon_h/\text{rad}^2, \theta_0 = 1.8326 \text{ rad}, \phi_\alpha = 1.0 \text{ rad}, \phi_\beta = \pi \text{ rad}$. A 和 B 的取值依据氨基酸的二级结构属性,对 α 螺旋: $A = 6\varepsilon_h, B = 5.6\varepsilon_h$; β 折叠: $A = 5.6\varepsilon_h, B = 6\varepsilon_h$; 其他情况: $A = B = 0$. 在计算非键能时,不同残基间相互作用简化为 NHP 3 种残基间的相互作用,具体参数 ε 和 Λ 见表 2. 从表 2 中可以看到,只有在疏水基团之间会产生吸引相互作用. 模拟结果都采用约化单位,即约化能量 $E^* = E/\varepsilon$, 约化时间 $t^* = t \sqrt{\varepsilon/m\sigma^2}$, 约化温度 $T^* = K_b T/\varepsilon$, 其中 K_b 为玻尔兹曼常数.

本文模拟采用真空中的分子动力学方法,试验性模拟实验中发现,蛋白质链经过 3×10^4 步长 (1.5 ns) 后基本能达到平衡状态. 各温度下蛋白质链的初始构型由单链蛋白质分子在高温 $T^* = 1.5$ 时随机产生,进行能量最小化后,降温到设定的环境温度,退火步长 $\Delta T^* = 0.05$ 并在每一个温度下保持 1.5 ns. 我们用 Nose-Hoover 方法来保持体系温度^[26,27]. 积分和松弛时间分别取 2.5 fs 和 0.025 ps,截断半径 $r_{\text{cutoff}} = 0.95 \text{ nm}$. 在我们的模拟中,选取的温度范围 T^* 从 0.1 到 1.0,总步长 1×10^7 即 500 ns,首先让体系构型在设定的温度下经过 2.5 ns (5×10^4 步长),所有的模拟结果是对后 497.5 ns 取平均值.

Table 2 Parameters ε and Λ in the nonlocal interaction term of potential function depending upon the nature of the two residues involved

	H	P	N
H	$\varepsilon = \varepsilon_h; \Lambda = 1$	$\varepsilon = 7/12\varepsilon_h; \Lambda = 0$	$\varepsilon = 2/3\varepsilon_h; \Lambda = 0$
P	$\varepsilon = 7/12\varepsilon_h; \Lambda = 0$	$\varepsilon = 1/6\varepsilon_h; \Lambda = -1$	$\varepsilon = 1/4\varepsilon_h; \Lambda = 0$
N	$\varepsilon = 2/3\varepsilon_h; \Lambda = 0$	$\varepsilon = 1/4\varepsilon_h; \Lambda = 0$	$\varepsilon = 1/3\varepsilon_h; \Lambda = 0$

2 结果和讨论

以一条典型的全 α 型蛋白质 1FCS 为例,从表 1 蛋白质链的基本信息可知该蛋白质链由 154 个氨基酸残基链接而成,其中中性、疏水和亲水型氨基酸个数分别为 43,50 和 61;此外,121 个氨基酸残基参与形成二级螺旋结构,在此蛋白质链的二级结构中不存在 β 折叠类型. 图 2 (a) 给出 1FCS 蛋白质体系势能随温度的变化趋势. 从图中可以得到如下结论:(1)键的伸缩能和弯曲能量都大于 0,这两部分势能随温度的升高呈线性增加,并在大部分温度下保持一致,仅在极高温下,键的伸缩能略高于键的弯曲能.(2)键的扭转

能是这四部分体系势能中最低的,从公式(3)中容易得出该部分能量总是小于 0. 在温度区间 $0.1 < T^* < 0.3$,扭转能随温度的升高而增加;在温度区间 $0.3 < T^* < 1.0$,扭转能随着温度的升高而降低, $T^* = 0.3$ 是体系扭转能发生转折的临界温度. (3)非键能在任意温度下都小于 0,说明在疏水——疏水氨基酸残基间发生相互作用对势能的贡献远大于其他类型氨基酸残基之间的相互作用. 非键能随温度的变化趋势和键扭转能正好相反,随约化温度升高先降低 ($0.1 < T^* < 0.2$),再增加 ($0.2 < T^* < 0.6$),然后趋于一个稳定值 ($0.6 < T^* < 1.0$). (4)蛋白质链体系总势能曲线的变化趋势,是各部分能量竞争的结果. 它的变化趋势

与非键能的变化趋势比较一致,在 $0.1 < T^* < 0.2$ 区间,总势能随温度增加而降低,可得在该温度区间,非键能随温度增加而降低的趋势占主导作用;当 $0.2 < T^* < 0.6$,势能随温度上升明显增加,在该温度范围,非键能以及键的伸缩和弯曲能量作用之和大于键扭转能随温度而减少的变化趋势;当 $0.6 < T^* < 1.0$,总的势能随温度的升高而缓慢增加,这时总势能的增加主要依赖于键的伸缩和弯曲能量.

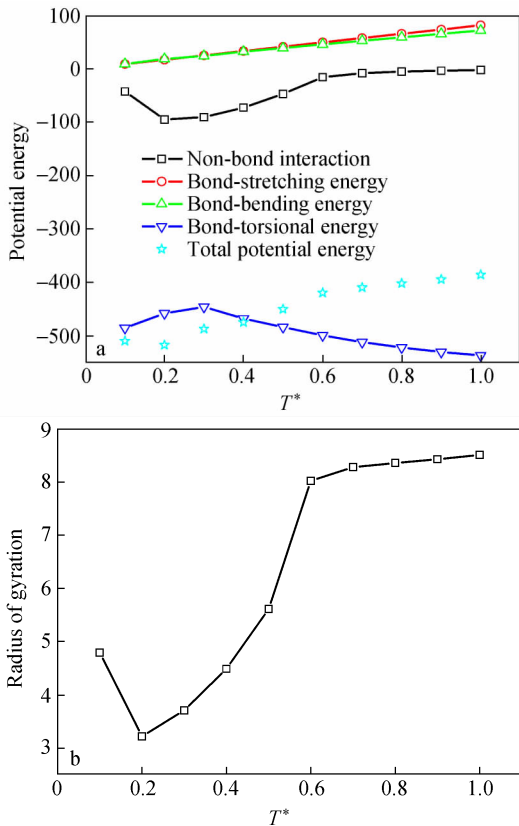


Fig. 2 The total and components of reduced potential energy (a) and the radius of gyration (b) versus reduced temperature T^* of 1FCS protein chain

通过上文的公式(1)、(2)、(3)可以看出随着约化温度的上升,蛋白质链中的化学键键长震荡更为剧烈,即 $(r - \sigma)^2$ 值增大,伸缩能增加;键角偏离平衡值的幅度增加,即 $(\theta - \theta_0)^2$ 值增大,弯曲能也随之上升;随着键长的伸长,弯曲程度的增加,使得蛋白质分子的整体空间结构更为疏散,非相邻氨基酸之间的距离增加,约化温度 $0.1 < T^* < 0.2$ 时,LJ 势能中,吸引项 r^{-6} 增长幅度快于排斥项 r^{-12} 促使非键能下降,当 $0.2 < T^* < 1.0$, r^{-12} 增长幅度快于 r^{-6} ,促使非键能上升.由此,大致可以得出该蛋白质链在温度 $T^* = 0.2$ 附近发生

相变.能量随温度变化的总体趋势受非键能影响较大,但同时又是键能(特别是键扭转能)和非键能的竞争结果.这一现象在任何一条全 α 型蛋白质的模拟结果中普遍存在.

同时,从图 2(b),1FCS 蛋白质链的回转半径随约化温度的变化趋势中也可以明显地观察到结构的变化.温度范围在 $0.6 < T^* < 1.0$,回转半径随温度的下降而缓慢下降,这主要源于键长的收缩;分子链的总体结构还是一个无规线团,在 $0.2 < T^* < 0.6$ 区间,回转半径随温度的下降而迅速下降,形成一个紧密的分子团;随着温度的进一步降低, $0.1 < T^* < 0.2$,分子链的回转半径反而随温度降低而增加,这说明蛋白质链形成一个取向结构,发生了折叠, $T^* = 0.2$ 也是图 2(a)中非键能和总的体系能量的转变温度.因此体系能量的研究结论对结构的预测提供了理论依据.

为了寻找更一般的规律,我们对 50 条不同链长的蛋白质体系进行研究.外界温度的变化范围 $0.1 \leq T^* \leq 1.0$.图 3 和图 4 分别为键伸缩能、键弯曲能与蛋白质链长的变化关系.在图中我们发现这些键伸缩能和键弯曲能与链长呈现较好的线性关系,并且斜率随温度上升而增加.比较图 3 和图 4 发现键伸缩能比键弯曲能的线性关系更好,特别在 $T^* = 0.2$ 时,图 3 中的数据点更接近拟合直线,它的线性拟合相关系数亦比图 4 中 $T^* = 0.2$ 的情形更接近 1.说明在低温下,蛋白质链的拉伸(由相邻 2 个氨基酸残基实现)比弯曲更容易实现(由相邻 3 个氨基酸残基实现).结合公式(1)、(2)分析,当约化温度一定时,在热力学平衡状态下,化学键键长震荡幅度以及化学键键角的变化幅度是一定的,即 $(r - \sigma)^2$ 及 $(\theta - \theta_0)^2$ 几乎不变,随着蛋白质链长的增加,化学键的总数增大,对它们的能量都是线性求和,因此与链长保持线性关系.当约化温度上升时,蛋白质分子震荡更为剧烈, $(r - \sigma)^2$ 及 $(\theta - \theta_0)^2$ 值增大,导致线性曲线的斜率增大,曲线更陡.

通过以上的分析,可以得出图 3 和图 4 中线性拟合直线的斜率即为蛋白质体系单个键在该温度下具有的键伸缩和键弯曲能.图 5 给出不同温度下单个键的伸缩和弯曲能量.从图 5 中可以看出,在约化温度 $0.3 < T^* < 1.0$ 范围内,单个键的伸缩能高于该温度下单个键的弯曲能量;在温度低于 0.3 时,则恰恰相反,弯曲能要高于伸缩能,这和我们前面讨论的结果相符合,即在低温状态

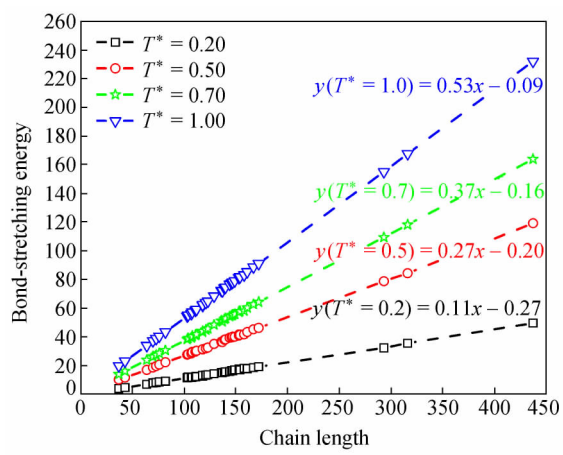


Fig. 3 Bond-stretching energy versus chain length of 50 all- α proteins

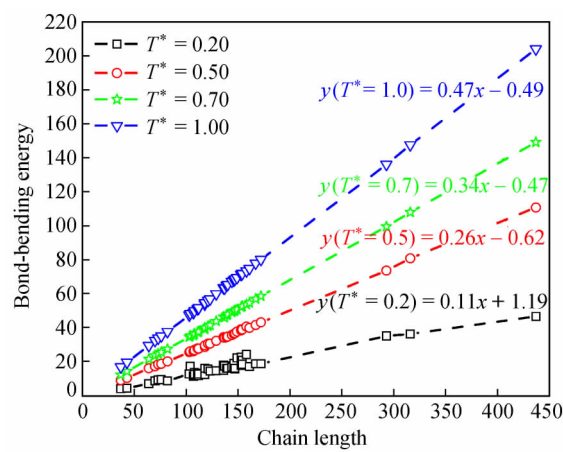


Fig. 4 Bond-bending energy versus chain length of 50 all- α proteins

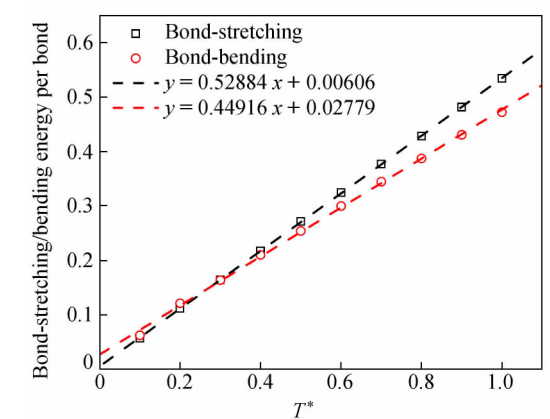


Fig. 5 Bond-stretching ($\square$) and bond-bending ($\circ$) energy of each bond versus reduced temperature T^*

下,蛋白质链的伸缩比弯曲更容易实现. 另外,在图中,我们对数据进行线性拟合也得到很好的结果,说明键的伸缩和弯曲幅度,即 $(r - \sigma)^2$ 及

$(\theta - \theta_0)^2$ 随温度是线性增加的,并且单个键的伸缩能随温度的增加要比单个键的弯曲能随温度的增加更为明显. 在图 1 中,蛋白质 1FCS 的键伸缩和弯曲能量也正体现了这种趋势.

图 6(a) 给出了不同温度下键扭转能的大小. 键扭转能的计算需要考虑蛋白质链的二级结构. 对全 α 类型的蛋白质,其二级结构大部分为 α -helix 螺旋结构(见表 1). 图中纵坐标表示键扭转能,横坐标为每条链中二级结构为 α 螺旋的数目. 从图中也发现较好的线性关系,键扭转能随 α 螺旋数目的增加而线性递减. 但是斜率随温度的变化不如键伸缩、键弯曲能明显. 在全 α 型蛋白质中,二级拓扑结构绝大部分由 α 螺旋构成, β 折叠所占比例非常小,即在公式(3)中的第二项基本为 0,蛋白质链的扭转能是对所有 α 螺旋结构对体系能量贡献的求和. 若与图 3、图 4 一样选择链长为横坐标,这样的线性关系则不成立,原因在于除了 α 螺旋以外的其他二级结构,如 β 转角的比例较高,有的甚至超过 α 螺旋的数目(如 1CCR, 2CDV 等),因此不能被忽略. 在图 6(b) 中,我们仅选择二级结构不包含 β 折叠的 38 条全 α 型蛋白质,键扭转能与 α 螺旋数目的线性关系则更为明显.

我们对二级结构不包含 β 折叠的 38 条全 α 型蛋白质的扭转能除以组成该条链二级结构为 α 螺旋的残基数目,并对样本进行统计平均,得到单个键的扭转能. 图 7 即为单个键扭转能随温度的变化关系曲线. 与图 2(a) 蛋白质 1FCS 体系的扭转能量随温度的变化曲线规律相同,在温度区间 $0.1 < T^* < 0.3$,扭转能随温度的升高而增加;在温度区间 $0.3 < T^* < 1.0$,扭转能随着温度的升高而降低, $T^* = 0.3$ 是扭转能发生转折的临界温度. 在 高分子链中,扭转角由连续的 3 个键的空间位置所决定,它随温度的变化,不再是简单的线性关系.

最后,我们对非键相互作用能进行讨论. 在我们采用的 NHP 模型中计算非键相互作用时要考虑氨基酸的亲疏水性质,这里只有疏水—疏水基团之间会产生吸引相互作用. 因此图 8 和图 9 中我们选择疏水残基 H 在蛋白质链中的比例作为横坐标. 图 8 分别给出在 $T^* = 0.1, 0.2, 0.5, 0.7$ 和 1.0 五个不同的环境温度下,50 条全 α 型蛋白质蛋白质体系的非键作用势. 从图可得,在温度较高时,非键相互作用能随着疏水性氨基酸在蛋白

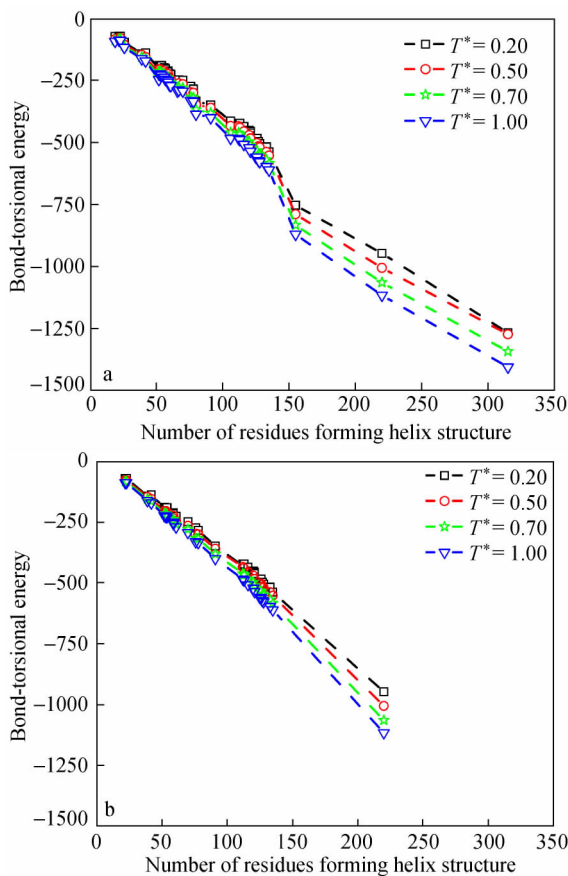


Fig. 6 Bond-torsional energy versus number of residues forming helix structure of 50 all- α proteins (a) or 38 all- α proteins without any β -sheet secondary structure (b)

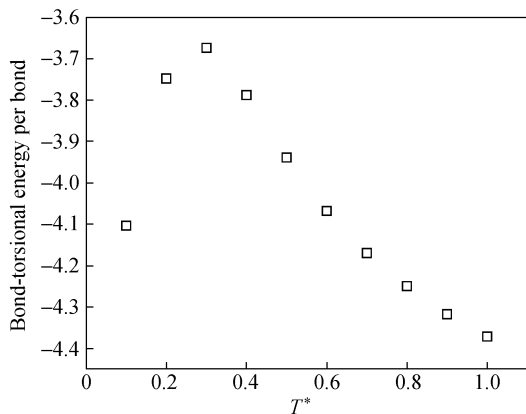


Fig. 7 Bond-torsional energy of each bond versus reduced temperature T^*

质链中所占比例的提高,能量降低,并呈现一定的线性趋势,说明 H 基团对稳定蛋白质起到积极的作用. 随温度升高,线性相关系数的绝对值越接近 1. 另外,绝大部分的蛋白质在 $T^* > 0.5$, 非键相互作用势能随着温度的升高而升高,但是在低温条件下与此变化趋势相反,除了 2CDV, 2CY3 和

2CTS, 绝大多数蛋白质在 $T^* = 0.2$ 时非键能要小于 $T^* = 0.1$ 时的非键能.

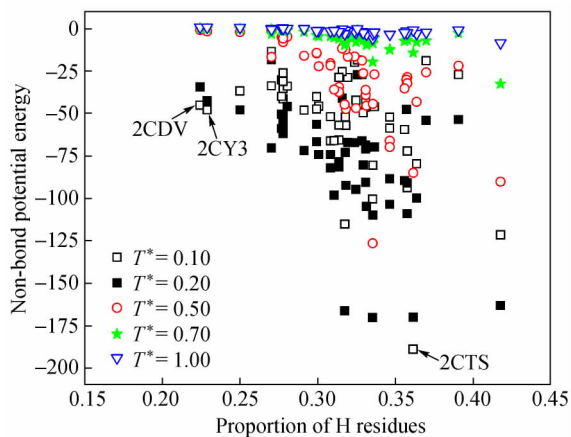


Fig. 8 Non-bond potential energy versus proportion of H residues in 50 all- α proteins

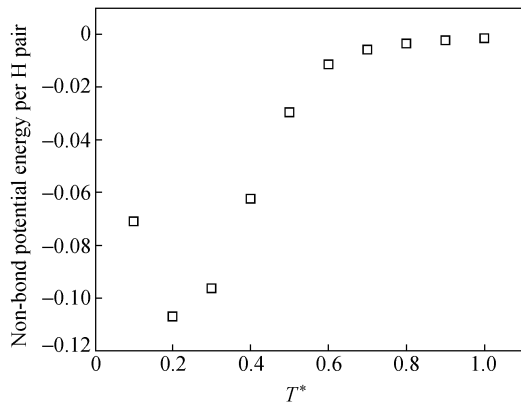


Fig. 9 Non-bond potential energy of each H—H pair versus reduced temperature T^*

考虑到非键作用成对出现,我们把体系非键作用势能除以可能的 H—H 相互作用对数目 $N_H \times (N_H - 1)/2$, 并对 50 条蛋白质进行统计平均, 给出不同温度下单个 H—H 对的非键能随温度的变化趋势图 9. 从图中我们可以得出这样的结论, 对于大多数蛋白质链, 非键能随温度的变化趋势与图 2(a) 一致, 即当约化温度 $0.1 < T^* < 0.2$ 时, 单个 H—H 对的非键能随温度的升高而降低, 在 $0.2 < T^* < 0.6$ 区间, 非键能随温度的升高而增加, 然后趋于一个稳定值 ($0.6 < T^* < 1.0$). 事实上, 蛋白质体系的非键相互作用还必须考虑其他两类残基 N 和 P 之间的两两相互作用. 但在我们的模型中, H—H 相互作用强度要远高于其他类型氨基酸残基间的相互作用, 讨论非键能随温度的变化趋势时, 忽略了其他类型的相互作用对. 图

8 中,2CDV,2CY3 出现的和绝大多数蛋白质链规律不一致的原因部分在于组成这 2 条蛋白质的疏水型氨基酸比例相对较小,一旦忽略了其他类型氨基酸的相互作用会造成相对大的误差.

通过键伸缩能、键弯曲能、键扭转能和非键相互作用能的分析,可以看出蛋白质在不同温度下的结构是由这 4 部分能量共同决定的.除了键伸缩能和键弯曲能与温度呈简单的线性关系,键扭转能及非键相互作用还和组成蛋白质的氨基酸疏水性、二级结构类型密切相关.从整条链的总势能随温度的变化趋势分析,它与非键能的变化趋势较为一致,因此非键能对蛋白质形成球状稳定结构起到关键作用,这个结论对一般的高分子也同样适用^[28].

3 结论

本文对蛋白质氨基酸的序列结构进行简化,建立 NHP 模型,结合势能函数表达式,根据分子动力学原理进行计算机模拟,得到不同的全 α 类型蛋白质的势能在不同温度下的变化情况.对模拟结果进行数值分析得到如下结论:(1)在任何

温度下,键的伸缩和弯曲能量都大于 0,且都随温度的升高而线性增加,其中键伸缩能比弯曲能随温度的增加幅度更为明显.当约化温度一定时,体系键伸缩能和弯曲能随蛋白质链长的增长而线性增加.(2)蛋白质链的扭转能随二级结构 α 螺旋的数目增加而线性下降,约化温度 $T^* = 0.3$ 是扭转能发生转折的临界温度,在临界温度以上,扭转能随温度升高而降低,在临界温度以下,扭转能随温度的降低而降低.(3)非键相互作用能的大小随疏水残基 H 在蛋白质组分中的比例增加而降低,并在高温状态呈现一定的线性关系,但此现象在低温状态下涨落很大,线性关系不明显.非键相互作用势在约化温度 $T^* = 0.2$ 达到最小值.(4)蛋白质链结构是由多种势能协同作用的结果,键能和非键能的相互竞争以及组成蛋白质链的氨基酸亲疏水性和二级结构的组分,都会影响蛋白质体系的势能.这些结论可以帮助我们找到最低能量下的天然构象,为蛋白质二级结构的确定,三级结构的预测打下良好基础,进而对蛋白质分子功能分析、性能改进乃至药物设计、疾病治疗提供理论依据.

REFERENCES

- 1 Levitt M, Chothia C. *Nature*, 1976, 261: 552 ~ 558
- 2 Marks D S, Hopf T A, Sander C. *Nat Biotechnol*, 2012, 30: 1072 ~ 1080
- 3 Furman O S, Wang C, Bradley P, Misura K, Baker D. *Science*, 2005, 310: 638 ~ 642
- 4 Kifer I, Nussinov R, Wolfson H J. *Proteins*, 2011, 79: 1759 ~ 1773
- 5 Murzin A G, Brenner S E, Hubbard T, Chothia C. *J Mol Biol*, 1995, 247: 536 ~ 540
- 6 Tong Huanping(全焕平), Zhang Linxi(章林溪). *Acta Phys Sin(物理学报)*, 2012, (5): 058701-1 ~ 6
- 7 Fromme P, Spence J C. *Curr Opin Struc Biol*, 2011, 24: 509 ~ 516
- 8 Barrientos L G, Dolan C, Gronenborn A M. *J Biomol NMR* 2000, 16: 329 ~ 337
- 9 Frans A A, Mulder F A, Filatov M. *Chem Soc Rev*, 2010, 39: 578 ~ 590
- 10 Wang Y J, Jardetzky O. *Protein Sci*, 2002, 11: 852 ~ 861
- 11 Klepeis J L, Lindorff-Larsen K, Dror R O, Shaw D E. *Curr Opin Struc Biol*, 2009, 19: 120 ~ 127
- 12 Tian J H, Garcia A E. *J Am Chem Soc*, 2011, 133: 15157 ~ 15164
- 13 Xu D, Zhang J, Roy A, Zhang Y. *Proteins*, 2011, 79: 147 ~ 160
- 14 Berger B, Leighton T. *J Comput Biol*, 1998, 5: 27 ~ 40
- 15 Fukugita M, Lancaster D, Mitchard M G. *Proc Natl Acad Sci USA*, 1993, 90: 6365 ~ 6368
- 16 Duan Y, Kollman P A. *Science*, 1998, 282: 740 ~ 744
- 17 Wang W, Xu W X, Levy Y, Trizac E, Wolynes P G. *Proc Natl Acad Sci USA*, 2009, 106: 5517 ~ 5522
- 18 Jiang Zhouting(姜舟婷), Zhang Linxi(章林溪), Chen Jin(陈进), Xia Agen(夏阿根), Ye Gaoxiang(叶高翔), Zhao Delu(赵得禄). *Acta Polymerica Sinica(高分子学报)*, 2003, (2): 207 ~ 210
- 19 Anfinsen C B. *Science*, 1973, 181: 223 ~ 230
- 20 Miyazawa S, Jernigan R L. *J Mol Biol*, 1996, 256: 623 ~ 644
- 21 Jiang Z T, Zhang L X, Sun T T, Wu T Q. *Chinese Physics B*, 2009, 18: 4580 ~ 4590
- 22 Jiang Z T, Zhang L X, Chen J, Xia A G, Zhao D L. *Polymer*, 2002, 43: 6037 ~ 6047
- 23 Jiang Z T, Zhang L X, Chen J, Xia A G, Zhao D L. *Polymer*, 2004, 45: 609 ~ 621

24 Chen Jin(陈进), Dong Zhijun(董志军), Zhang Linxi(章林溪), Jiang Zhouting(姜舟婷), Zhao Delu(赵得禄). Acta Polym Sinica(高分子学报), 2003, (4): 577 ~ 580

25 Sangha A K, Keyes T. J Phys Chem B, 2010, 114: 16908 ~ 16917

26 Nose S. J Chem Phys, 1984, 81: 511 ~ 519

27 Hoover W G. Phys Rev A, 1985, 31: 1695 ~ 1697

28 Jiang Z T, Xu P, Sun T T. Chinese J Polym Sci, 2012, 30: 45 ~ 55

A Study on the Energy Transition of All- α Proteins by Molecular Dynamics Simulation

Zhou-ting Jiang^{1*}, Ting-ting Sun², Ya-nan Wang¹, Yu-dong Chen¹, Zhi-xing He¹, Peng Xu¹

(¹ College of Science, China Jiliang University, Hangzhou 310018)

(² School of Information and Electronic Engineering, Zhejiang Gongshang University, Hangzhou 310018)

Abstract All- α proteins simplified by NHP model with its secondary structure information were studied by molecular dynamics simulation. The transitions of potential energy in 50 all- α proteins under different temperatures were well discussed in this paper. It is concluded that the bond-stretching and bending energies increase linearly with the chain length. The bond-stretching and bending energies of each bond also increase linearly with the temperature due to more obvious bond vibration under high temperatures. The slope of the linear relationship between bond-stretching energy and temperature is larger than the one of bond-bending energy as the function of T^* . The bond-torsional energy decreases linearly with the number of α -helix in each protein chain, instead of the chain length. It increases with the temperature when the temperature is lower than 0.3. And then it decreases with the temperature increasing. Non-bond energy is mainly dependent on hydrophobic-hydrophobic interactions. It doesn't show obvious linear relationship due to the mutual interactions between hydrophilic, hydrophobic, or neutral residues. The transition temperatures are shown in bond-torsional and non-bond potential energies. The native structure is related to the mutual effect of bond and non-bond potential energies, which compete with each other for lower potential energy. These conclusions provide a natural realization of understanding the structure of proteins.

Keywords All- α proteins, Molecular dynamics simulation, Energy transition

* Corresponding author; Zhou-ting Jiang, E-mail: z. jiang@ cjlu. edu. cn