

目 录

1 绪论	1
1.1 生皮的基本概述	1
1.1.1 生皮蛋白质的组成	2
1.1.2 生皮蛋白质的分类及结构	3
1.1.3 生皮蛋白质的性质及应用	6
1.2 皮革的热稳定性	8
1.2.1 皮革热稳定性的重要性	8
1.2.2 皮革热稳定的研究现状	9
1.3 本课题的研究内容及意义	11
2 分子动力学	13
2.1 基本原理	13
2.2 运动方程的求解	14
2.3 势函数	15
2.4 一些重要术语	15
2.5 分子动力学的应用	16
2.6 本章小结	17
3 实验部分	18
3.1 实验所用软件	18
3.1.1 TINKER 软件	18
3.1.2 Force Field Explorer 软件	19
3.1.3 MATLAB 软件	21
3.2 实验前期准备	22
3.2.1 PDB 文件的准备	22
3.2.2 Key 文件的准备	23
3.3 实验过程	25
3.4 模拟轨迹的分析	28
3.4.1 模拟轨迹的分析方法	29
3.4.2 模拟轨迹的分析过程	31
3.5 本章小结	33
4 实验结果的分析与讨论	34

4.1 RMSD 分析.....	34
4.2 轨迹取样分析.....	35
4.2.1 260K 轨迹取样分析	36
4.2.2 300K 轨迹取样分析	37
4.2.3 340K 轨迹取样分析	38
4.2.4 380K 轨迹取样分析	39
4.3 Rg 分析	41
4.4 氢键分析.....	42
4.5 本章小结.....	45
致 谢.....	46
参考文献.....	47
附 录.....	51

1 绪论

天然皮革之所以千百年来经久不衰，是因为它具有很多优良的性能。首先动物皮跟人的皮肤一样主要是由胶原蛋白组成，所以天然皮革制成的服装即使贴身穿也不会有其他化学纤维所带来的不良反应，穿着柔软舒适；其次天然皮革是由天然蛋白质纤维编织而成的，所形成的多孔性使得它具有良好的透气性，而且蛋白质分子链上的亲水基团赋予了天然皮革良好的吸湿性；再次天然皮革拥有独特的纹理效果及坚固耐穿的特性。虽然具有以上种种的优良性能，天然皮革还是有很多不足需要进一步的改进，例如天然皮革无论是在保管，加工还是后期的使用过程中都会受到各种热因素的影响，比如制革过程中的湿热处理，制鞋过程中的热定型等，所以对皮革热稳定的研究具有现实意义。

关于天然皮革热稳定性的研究有两种方法，宏观研究法和微观研究法。宏观研究法主要是通过大量的对比实验研究来确定皮革热稳定性能的好坏，这种方法存在一定的局限性和误差。我们主要是从微观方面来研究天然皮革的热稳定性。因此，我们首先来了解一下天然皮革的显微结构及其化学成分，以促进我们深入分析。

1.1 生皮的基本概述

生皮从外观上看可分为皮板和毛被两大部分。在显微镜下我们观察生皮的纵切面，可清楚的看到皮板分为3层：表皮、真皮和皮下层^{[1]6}，如图1.1。在制革过程中表皮层和皮下层都要被除去，只保留真皮层。

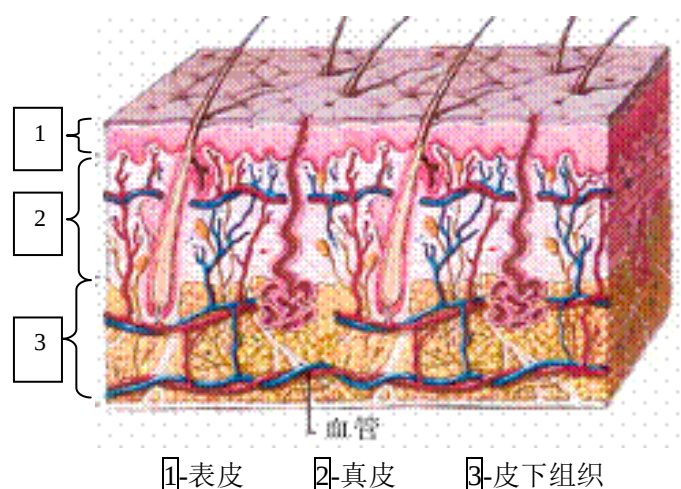


图 1.1 生皮的纵切面示意图

Fig.1.1 The longitudinal section of raw hides and skins

生皮的组成成分十分复杂，其中最主要的是蛋白质，还包含了水分和脂类等其它成分。表 1.1 列出了生皮各个组分的含量。

表 1.1 生皮的组分/ (%) ^{[1]10}

Table 1.1 The Component of Hides and Skins / (%)

组分	含量	组分	含量
蛋白质	30-50	无机盐	0.3-0.5
水分	60-75	碳水化合物	<2
脂类	2.5-3.0		

生皮所含有的蛋白质种类很多，其中表皮和毛发主要是由角蛋白组成，真皮的蛋白质主要是由胶原蛋白构成的纤维状胶原，含量达 80%~85%，另外还有弹性蛋白、网硬蛋白等，纤维间质中含有白蛋白、球蛋白和类粘蛋白等，它们的含量会随着动物的种类、性别、年龄和生活条件的不同而有所变化^{[1]10}。由此可见胶原蛋白是组成天然皮革的基本物质，要想研究天然皮革的热稳定性必须首先了解胶原蛋白。

1.1.1 生皮蛋白质的组成

生皮主要的蛋白质为胶原蛋白，它的英文“collagen”是从希腊文演化过来的。胶原蛋白是蛋白质的一种，其组成的元素主要是 C、H、O、N、S 五种，也有一些蛋白质中还含有 P，卤素元素和一些金属元素，例如 Fe、Cu、Zn、Cd 等，其中氮的含量较其他蛋白质要高^{[1]11}。

蛋白质借助酸、碱或者酶的作用会逐渐发生水解，得到最终产物—L 型的 α 氨基酸，也就是蛋白质的基本单位。一般的蛋白质含有 20 种氨基酸，表 1.2 列出了几种常见动物皮胶原的氨基酸组成。

表 1.2 几种常见动物皮胶原的氨基酸组成/ (残基个数/1000 个残基) ^[2]

Table 1.2 The amino acid composition in several common animals' collagen

蛋白质 材料来源	小牛皮胶原	公牛皮胶原	猪皮明胶
丙氨酸	112	108	112
甘氨酸	324	334	328
缬氨酸	20	19	22
亮氨酸	25	25	24
异亮氨酸	11	11	10
脯氨酸	138	129	130

表 1.2 (续) 几种常见动物皮胶原的氨基酸组成/(残基个数/1000 个残基)
Table1.2 (continue) The amino acid composition in several common animals' collagen

蛋白质 材料来源	小牛皮胶原	公牛皮胶原	猪皮明胶
苯丙氨酸	3	13	14
络氨酸	3	5	3
丝氨酸	36	38	37
苏氨酸	18	17	18
胱氨酸	—	—	—
蛋氨酸	5	7	5
精氨酸	50	48	48
组氨酸	6	5	6
赖氨酸	27	27	27
天冬氨酸	46	48	46
谷氨酸	72	72	74
羟脯氨酸	94	94	96
色氨酸	—	—	—

从表 1.2 中我们可以看出胶原蛋白氨基酸组成的几大特点^{[3]4-5}:

(1) 胶原蛋白中不含胱氨酸和色氨酸,但也有人认为胶原蛋白其实并不是不含这两种氨基酸,只是这两种氨基酸的含量很少而已;

(2) 甘氨酸的含量很高,三种胶原蛋白中甘氨酸的含量分别为 32.4%, 33.4%和 32.8%,几乎占了整个胶原蛋白氨基酸含量的 1/3;

(3) 胶原蛋白中比较特别的氨基酸为赖氨酸和羟脯氨酸,因为在其他蛋白质中不存在或很少存在它们;

(4) 胶原蛋白中脯氨酸和羟基脯氨酸的含量比较高,在三种胶原蛋白中它们的含量加起来都超过了 20%,与其他种类的蛋白质相比,它们的含量是最高的。

1.1.2 生皮蛋白质的分类及结构

胶原蛋白是一种细胞外基质的结构蛋白,随着生物学的发展,人们已经认识到胶原蛋白不仅仅是一个蛋白质的总称,而是富有多多样性和组织分布特样性,并与各组织和器官功能密切相关的一种功能蛋白。到目前为止,人类已经发现了 27 种不同类型的胶原蛋白^[4]。表 1.3 列举了 21 种不同类型胶原蛋白的分子组成及组织分布情况。

表 1.3 不同类型胶原蛋白的分子组成及组织分布^[5]

Table 1.3 Composition and distribution of various collagen types

类型	分子组成	主要分布组织
I	$[\alpha_1(\text{I})]_2\alpha_2(\text{I})$	皮肤、肌腱、韧带、骨骼等
II	$[\alpha_1(\text{II})]_3$	软骨、玻璃体、髓核、椎间盘
III	$[\alpha_1(\text{III})]_3$	皮肤、血管、肌肉
IV	$[\alpha_1(\text{IV})]_2\alpha_2(\text{IV})$	基底膜
V	$\alpha_1(\text{V}), \alpha_2(\text{V}), \alpha_3(\text{V})$	肺、角膜、骨骼
VI	$\alpha_1(\text{VI}), \alpha_2(\text{VI}), \alpha_3(\text{VI})$	真皮、软骨、胎盘、肺、静脉血管壁、椎间盘
VII	$[\alpha_1(\text{VII})]_3$	皮肤、口腔粘膜、子宫颈
VIII	$[\alpha_1(\text{VIII})]_2\alpha_2(\text{VIII})$	内皮细胞等
IX	$\alpha_1(\text{IX}), \alpha_2(\text{IX}), \alpha_3(\text{IX})$	软骨、玻璃体、角膜
X	$[\alpha_1(\text{X})]_3$	增生软骨
XI	$\alpha_1(\text{XI}), \alpha_2(\text{XI}), \alpha_3(\text{XI})$	软骨、玻璃体
XII	$[\alpha_1(\text{XII})]_3$	软骨膜、韧带、腱
XIII	$[\alpha_1(\text{XIII})]_3$	表皮、毛囊、肌内膜、肠、肺、肝
XIV	$[\alpha_1(\text{XIV})]_3$	真皮、腱、静脉血管壁、胎盘、肺、肝
XV	$[\alpha_1(\text{XV})]_3$	纤维原细胞、平滑肌肉细胞、肾、胰腺
XVI	$[\alpha_1(\text{XVI})]_3$	纤维原细胞、胞衣、角化细胞
XVII	$[\alpha_1(\text{XVII})]_3$	表皮
XVIII	$[\alpha_1(\text{XVIII})]_3$	肝及胃细胞
XIX	$[\alpha_1(\text{XIX})]_3$	横纹肌肉瘤
XX	$[\alpha_1(\text{XX})]_3$	角膜上皮细胞、胚胎皮肤、胸骨软骨、腱
XXI	$[\alpha(\text{XXI})]_3$	静脉血管壁

以上21种不同类型的胶原蛋白的区别在于他们结构的复杂性和多样性，其中分子中额外的非螺旋域的分布及范围的不同是它们之间最主要的区别，这个区别也决定了蛋白质的特性。胶原蛋白按照功能可以分为成纤维胶原和非成纤维胶原。I、II、III、V、和XI型为成纤维胶原，他们是胶原蛋白家族中数目最多分布最广的一类蛋白，占胶原家族的90%，其余为非成纤维胶原。这两组胶原蛋白最主要的区别在于组成胶原域的三股螺旋结构是连续的还是非连续的，成纤维胶原中的三股螺旋结构是长而连续的，非成纤维胶原中的三股螺旋结构至少有一处是中断的。

自从英国的伯纳尔在蛋白质结构的研究中发现，有些蛋白质具有比三级结构更加复杂的结构，他把这种更复杂的结构称为四级结构。目前这种关于四级结构的说法已经被各国科学家所认同，图1.2为蛋白质的四级结构图。胶原蛋白作为蛋白质的一种，也具有四级结构。

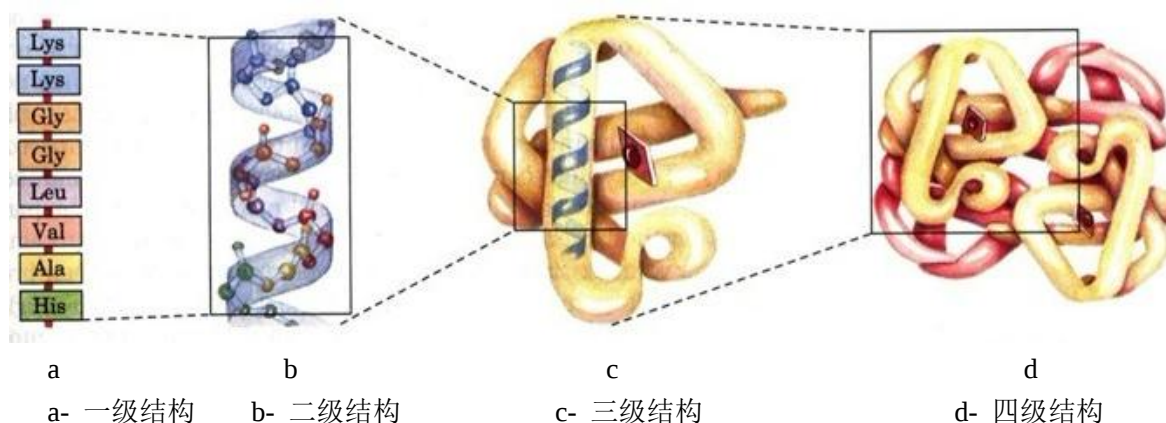


图1.2 蛋白质的结构

Fig.1.2 The structure of protein

蛋白质的一级结构（primary structure）是蛋白质最基本的结构，是指蛋白质多肽链中氨基酸残基的排列顺序，由基因上遗传密码的排列顺序所决定。蛋白质的一级结构告诉我们该蛋白质含有多少种氨基酸，每种氨基酸的含量是多少以及每个氨基酸之间是怎么连接成多肽链的，它决定了蛋白质的多级结构。胶原蛋白一级结构的主要特点是胶原域(胶原分子中的具有 α 链形成的三股螺旋区域)有着重复序列存在,形成了三肽Gly-X-Y的重复区域, Gly代表甘氨酸, X、Y通常是指脯氨酸或者羟脯氨酸, 这种重复区域对胶原蛋白的结构起着重要作用, 也说明了胶原蛋白中存在大量的甘氨酸, 大约占据1/3^{[3]6}。

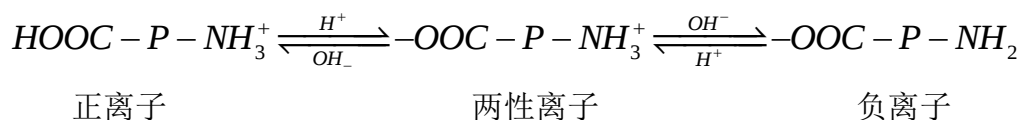
蛋白质的二级结构（secondary structure）是指多肽链主链的空间排列情况，二级结构是三级结构的结构单元，也是蛋白质空间构象的基础，所以也把它们称为构象单元。蛋白质二级结构一般包括四种： α -螺旋， β -弯曲， β -拐角和无规卷曲^[6]。胶原的二级结构是由三条肽链组成的三股非 α 左手螺旋通过甘氨酸残基的肽键之间形成的氢键交联在一起而形成的右手超螺旋结构^[7]。把胶原蛋白的螺旋称为非 α 螺旋的原因是胶原蛋白每圈螺旋包含的氨基酸残基数目，螺旋与螺旋之间的距离以及相邻残基间的间距这些数值都与标准 α 螺旋的要求不一致。

蛋白质的三级结构（tertiary structure）是指包含侧链在内的整个多肽链所形成的三维空间结构。蛋白质主要是靠氨基酸残基侧链的极性基团所产生的离子键、氢键、二硫键、范德华力以及疏水键等作用力来维持三级结构的稳定性^[6]。胶原蛋白的三级结构是指直径约1.5nm的胶原分子并行排列并通过共价交联形成了胶原微纤维时相互之间的三维空间关系^[8]。

蛋白质的四级结构（quaternary structure）是指由一些具有特殊三级结构的肽链通过次级键相互组合而形成的空间结构。胶原的四级结构一般指原胶原分子按首尾错位四分之一的方式有规则错位并平行排列形成稳定并且韧性很强的原纤维（fibril）^[9]。

1.1.3 生皮蛋白质的性质及应用

胶原蛋白是一种聚两性电解质，这主要包含两个方面的原因：①胶原的每个肽链有许多侧基，它们具有酸性或碱性；② α -氨基和 α -羧基位于胶原每个肽链的两端。这些基团都有接受或者给予质子的能力，因此在不同pH值的溶液中胶原会解离成为带正电荷或者负电荷的离子：



对于某一种胶原蛋白，当它在特定PH值的溶液中所带的正负电荷数相等时，也就是胶原蛋白既不带正电也不带负电，我们把此时的PH值称为这一蛋白质的等电点。牛皮的等电点为7.5-7.8，溶液为偏碱性，即牛皮胶原蛋白带负电荷，此时若对牛皮胶原蛋白进行电泳则蛋白质向正极移动^{[3]9}。不同蛋白质的等电点不同，偏酸或者偏碱，与该蛋白质酸性基团和碱性基团的数量有关，牛皮胶原蛋白的等电点偏碱性是由于牛皮胶原蛋白侧连上的碱性基团的数量比较酸性基团多的缘故。当蛋白质位于等电点的时候，它的溶解度、导电率、膨胀率等性质都处于最小值。

正是由于胶原中存在酸性基团和碱性基团所以胶原能够在溶液中与酸和碱结合，在这个结合过程中胶原由于离子交联键和氢键被打开而吸水发生膨胀，如果时间延长的话会导致胶原键的破坏从而使它发生胶解。在制革生产中我们会利用胶原在盐溶液中离子键被盐打开而吸水的这种性质。但是不同的中性盐对胶原蛋白的作用是不同的。例如，在原料皮防腐时加入氯化钠是为了使原料皮脱水以抑制微生物的生长；浸水时加入中性盐效应较强的物质，如多硫化钠，可以促进干皮尽快恢复到鲜皮状态；削皮时用硫酸钠处理，使之脱水，利于削匀操作；浸酸时加入氯化钠，可以避免胶原在酸中膨胀。

物理性质方面，胶原蛋白的空间结构在加热、紫外线照射、机械力等物理因素下或者加入酒精，强酸强碱等化学试剂时，会导致胶原蛋白结构的改变而引起物理化学和生物性质的改变，这种现象叫做胶原蛋白的变性。也就是说，当外界条件发生改变时，胶原蛋白内部维持它稳定结构的作用力，如氢键、范德华作用力、疏水作用和盐键等非共价键（一些弱的相互作用）遭到破坏，使得胶原蛋白三股螺旋结构转变为松散无序的结构，从而引起性质的改变。此外，由于胶原蛋白分子表面存在许多极性基团，如氨基、羧基、羟基、肽基等，这些极性基团会以氢键的方式与水分子结合使胶原发生膨胀，也就是胶原发生了水和作用^{[3]15}。

从前面关于胶原蛋白结构和性质的介绍我们知道胶原蛋白是由特殊氨基酸组成的，具有独特而稳定的三股螺旋结构，而结构又决定了性质，因此胶原蛋白拥有很多优良的性能，如良好的生物相容性、低抗原性和可降解性等生物活性，因此它的应用范围十分

广阔。

在食品方面,由于胶原蛋白所含氨基酸种类丰富,其中甘氨酸、脯氨酸和羟脯氨酸的含量也很高,而且胶原蛋白还包含其他蛋白质中不存在的羟基赖氨酸或者含量很少的羟脯氨酸、赖氨酸,因此胶原蛋白的营养十分丰富。在食品添加剂方面,Meullenet等人研究表明,添加胶原2%、水20%左右时,腊肠在外观、质地和口感等方面达到最佳效果^[10];胶原蛋白可作为冷冻食品的改良剂,在冰激凌中添加0.25% - 0.60%的明胶可以保持冰激凌细腻的口感同时减缓了它的融化速度^[11];在制作肠衣方面,由于胶原蛋白营养丰富并且在热处理过程中胶原与肉类食品的收缩率几乎一致,因此在美国、日本等发达国家已经普遍用胶原蛋白肠衣代替天然肠;在保健食品方面, Mattsson等人发现胶原蛋白能够减弱风湿性关节炎的发生^[12];李彦春等用小白鼠做实验发现,加入了胶原蛋白的钙制剂能改善骨质疏松,促进小白鼠的生长发育,说明胶原蛋白能促进钙质的吸收^[13]。

胶原蛋白与人体皮肤的结构类似,因而可以被人体皮肤快速吸收而达到修复损伤皮肤的目的;胶原蛋白中富含的氨基酸可以补充皮肤所需营养,促进皮肤组织的新陈代谢,达到滋润皮肤延缓衰老的效果;胶原蛋白中含有大量的保湿因子,甘氨酸、丙氨酸等,同时分子链上大量的羟基和氨基可以提高皮肤组织细胞的贮水能力,从而达到保湿的效果,因此胶原蛋白广泛的运用于化妆品^[14-15]。Youngh采用超低温萃取法应用纳米技术制成的新一代胶原蛋白产品-雅佗胶原蛋白分子,可使肌肤充满弹性,恢复细腻和活力^[16]。

由于胶原蛋白具有良好的生物材料性该具备的性能:胶原蛋白结构重复性大因此免疫性低,一般机体不对其产生排斥反应;胶原的四级结构使其具有凝聚能力,能与血小板一起形成血栓控制流血^[17];胶原蛋白具有良好的生物相容性;胶原蛋白的胶原肽键只能在胶原酶的作用下肽链才会被打断,螺旋结构才会被破坏,因此具有良好的可控生物降解性。王永胜等人研究发现海绵状胶原蛋白在短时间内可达到满意的止血效果,并且对伤口的愈合和恢复也具有独特的效果^[18]; Judith Hohlfeld等的研究表明利用生物工程技术培育出的婴儿皮肤组织可治愈不同程度的烧伤,而且新长出的皮肤也很少表现出肥大增生和抗性,因此在治愈烧伤方面具有不可低估的潜力^[19]。

胶原蛋白分子链上有很多极性基团,这些极性基团能与纸张中植物纤维分子中的羟基形成化学键,进而增大纸张的强度。付丽红等人用实验表明,当胶原蛋白含量为2%时纸张的裂断长和耐破度效果最佳,胶原蛋白含量为4%时耐折度最好,加胶原蛋白与未加胶原蛋白的纸张相比,变形温度提高了25.02K,所要吸收的热量提高了 $2.92 \times 104 \text{J/kg}$,热分解温度增加了1.86K,即加入胶原蛋白后纸张的热稳定性增加了^[20]。

胶原蛋白与皮革胶原结构相似,因此可以利用胶原蛋白对皮革的一些制剂进行改性从而使这些制剂渗透到皮革胶原的内部提高皮革的性能。贾鹏翔,汤克勇从胶原蛋白与皮革有很好的相容性这个角度出发,提出了用胶原蛋白改性丙烯酸类复鞣剂的方案,经实验表明:当过硫酸铵,马来酸酐,丙烯酸,胶原蛋白的添加量之比为0.25: 2: 3: 0.5

时形成的共聚物，当共聚反应温度为80℃，共聚反应时间为1h，胶原改性反应温度为60℃，改性反应时间为3h时所形成的改性类复鞣剂用于皮革复鞣工艺不仅可以使皮革的粒面光滑细腻，提高了皮革的强度，弹性和丰满度，而且相对于没有用胶原蛋白改性的丙烯酸类复鞣剂而言，改性后的复鞣剂能提高皮革的湿热稳定性和透湿率^[21]。

从前面关于生皮的介绍我们知道真皮中最主要的蛋白质是胶原蛋白，特别是I型胶原蛋白，而蛋白质的结构决定性能，因此要想了解皮革的热稳定性能就要了解胶原蛋白的结构，特别是I型胶原蛋白的结构，从结构出发找到与皮革热稳定性相关的因素。

1.2 皮革的热稳定性

1.2.1 皮革热稳定性的重要性

皮革作为一种材料，耐热性是评价其性能好坏的一项重要指标。在皮革的加工过程中，为了促进染料和油脂的溶解和分散，加快渗透的速度，同时加快染料和油脂与革的结合从而提高吸尽率，在染色和加脂工序中会对皮革进行50~60℃的湿热处理^[22]；在制鞋过程中，第三道工序-热定型非常重要，它的作用有两个：一是松弛由第二道工序绷邦操作所造成的鞋面不同部位不均匀的受力而使鞋面平整顺滑；二是使鞋面固定成楦的形状。第一个作用需要对皮革进行湿热处理而消除应力，第二个作用为了使鞋面固定成所需要的形状必须经过120℃~130℃的高温处理。在把鞋底和鞋跟固定的这一操作过程中皮革可能要经受高达170℃的高温处理，这个温度取决于固定方法是采用粘和，硫化还是注塑^[23]；在皮革产品的保养过程中，为了防止皮革变硬或者收缩，皮革也不允许曝晒或烘烤，即使是在熨烫的时候也要严格的控制熨烫温度和熨烫方法，熨斗不能直接接触皮革，也不可以在某一部位停留过长的时间。因此，无论是从皮革的加工过程，使用过程以及皮革产品的保养过程都可以看出，皮革必须具有一定的热稳定性，否则高温会引起皮革面积减少，体积缩小的现象，也就是发生收缩变形，即使收缩后仍然能够继续使用也大大价低了皮革的经济效应，而且如果温度过高的话皮革会被破坏而失去原有的使用价值。

汤克勇，贾鹏翔等人通过实验方法研究了皮革收缩时，它的相关性能发生变化的情况。力学性质方面，通过对比分析加脂后的皮革收缩前后的应力应变曲线（将皮革两端固定时皮内产生一种内力，单位面积皮革所受的这种内力称为皮的应力）及力学各参数的变化得出，收缩后加脂皮革的断裂伸长率以及抗张强度都大幅度提高，表现出良好的力学性质，但是收缩后皮革的弹性模量也变大，使得皮革的硬度提高而不利于穿着；孔隙率的变化方面，通过比较未加脂革，加不同加脂剂的加脂革，以及加相同加脂剂但加脂剂的剂量不同的加脂革在收缩前后孔隙率的变化得知，不管是否加脂或者加什么类型

的脂或者加多少剂量的脂，收缩后皮革的孔隙都会变小，孔隙率都会降低，这种变化就阻碍了皮革胶原蛋白分子上亲水基团对水分子的传递，因此皮革的透湿性及吸水性都明显降低^[24]。

综上所述，皮革的热稳定性是皮革非常重要的一项性能，无论是对加工还是生产都具有非常重要的意义。皮革的热稳定性能如果不好，那么在湿热或者干热处理时皮革就会发生收缩、变硬，不仅影响外观而且降低了皮革的舒适性和经济价值。皮革收缩后它的力学性质虽然有所提高，但是孔隙率的下降导致皮革吸水性和透湿率的降低使得皮革作为服装材料的一大优势也不复存在。因此需要对皮革的热稳定性进行研究，从而了解皮革热稳定机制，进一步指导皮革热稳定性的提高，这对皮革产品的使用性能和加工性能的提高具有重要的现实意义。

1.2.2 皮革热稳定的研究现状

Chahine指出影响胶原蛋白热稳定性的因素有很多，大致包括两大类：生物因素和非生物因素，生物因素主要是指环境、动物的生活条件、动物种类、动物的年龄以及动物身上不同的部位^{[25]53}。目前已经有研究表明胶原蛋白的收缩温度与胶原蛋白中的氨基酸含量、氨基酸的羟化程度以及氨基酸在分子链中的位置有关。例如Menashi等人测得鱼皮中的亚氨基酸含量很低，所以鱼皮的收缩温度低于哺乳动物，也就说明了胶原蛋白的收缩温度与亚氨基酸的含量相关^[26]。

非生物因素包括了加热介质的性质、所处的离子环境、鞣质情况、盐以及动物的老化程度^{[25]53}。这里主要介绍皮革热稳定性非生物因素方面的研究。

陈新江等人分别用水、水蒸气和甘油作为加热介质，测定三种加热介质对四种不同的革：铬鞣革、铝鞣革、植鞣铬、醛鞣革收缩温度的影响，结果表明用甘油，水作为加热介质进行测量时，革出现脱鞣现象，所测得的皮革收缩温度比较低，结果不合理，而用水蒸气作为加热介质时革的脱鞣现象不太明显，所测得的结果比较合理，能较真实的反应各种鞣制皮革在湿热环境下的稳定性^[27]。

Komanowsky研究认为通过改变离子环境，也就是改变介质的PH值使其偏离胶原蛋白的等电点，此时胶原蛋白会发生溶胀现象而增加了胶原蛋白分子链上活性基团间的距离，使得分子间的作用力降低而导致热稳定性的减弱^[28]。

鞣制是将原料皮变为革的重要步骤，它对胶原蛋白的热稳定性有着重要的影响。不同的鞣制剂与胶原蛋白的结合方式不同，鞣制机理也就不同，从而对胶原蛋白热稳定性的影响也就不一样。铬鞣制过程中，主要是铬鞣液中的铬配合物阳离子与胶原蛋白分子侧链上的极性基-羧基发生配位结合形成强的交联作用，所形成的交联键为配位键。植物鞣的机理则是鞣制酚羟基与胶原蛋白分子主链上的酰胺基形成氢键为主^[29]。醛鞣则主

要是甲醛单体与赖氨酸侧链上的 ϵ -氨基和精氨酸侧链上的胍基之间形成亚甲基桥，所形成的交联键为共价键^[30]。相同的鞣制剂在不同的溶剂中时对胶原蛋白的热稳定性的影响也是不同的。王亚娟等人采用不同极性的有机溶剂，分别处理不同的革样，通过测定其收缩温度得出结论：有机溶剂中甘油对各种革耐湿热稳定性的影响较大，浸泡了甘油的革与未浸泡的革相比收缩温度下降了 12.9°C ^[31]。Takenouchi等人在铬含量处于1-5%之间时，分别把甲酸盐、乙酸盐和草酸盐加入到CrC13.6H2O溶液中配置成蒙圈的铬鞣液进行皮粉的鞣制，发现三者的收缩温度分别为 $80\sim 100^{\circ}\text{C}$ ， $70\sim 100^{\circ}\text{C}$ ， 80°C 左右。并通过比较不同组分的甲酸合铬络合物所鞣制的皮粉的热稳定性得出+3价的络合物对皮粉热稳定的贡献比+2价络合物的要大，不同组分的乙酸合铬络合物所鞣制的皮粉中+2价和+3价阳离子络合物相对于阴离子和中性络合物而言，对皮粉热稳定的作用要大，一价草酸合铬络合物与二价草酸合铬络合物相比，前者对皮粉热稳定性的作用大^[32]。

康俊霞通过实验得出较低浓度的盐溶液中胶原蛋白的收缩温度降低，可能是由于 Na^+ 影响了胶原蛋白氨基酸残基之间的相互排斥力^[33]，与REGINA^[34]得出的结论一致。

已经有研究（如用差示扫描量热法，differential scanning calorimetry, 简称为DSC）证明了皮革的湿热收缩是一个焓增的过程^[35]。由于DSC方法能得到皮革收缩时的焓变，并且由于胶原变性时吸收的热量很多便于观察研究和分析，所以DSC方法比较适合研究胶原材料。

有一批学者从热力学和动力学的角度研究了皮革的收缩。如周鲁等人根据Gibbs-Helmholtz公式得到皮革的热收缩温度 T_s 为：

$$T_s = \Delta H / \Delta S \quad \text{式 (1-1)}$$

式中： ΔH - 焓变； ΔS - 熵变

依据皮革的是热收缩是一个焓增的过程，再根据热力学原理推导出皮革热收缩不仅是一个焓增的过程，同时也是一个熵增的过程。由公式（1-1）可以看出，皮革的热收缩温度 T_s 与 ΔH 和 ΔS 同时有关，既然皮革的热收缩过程是焓跟熵都增加的过程，那么二者增加的多少就对热收缩的温度有影响，要想提高皮革的热稳定性也就是要提高皮革的热收缩温度就要求我们要控制吸热过程中的焓变和熵变，尽量增加焓变量而减少熵变量；皮革在湿热条件下，环境和皮革之间存在热量和水分传递，此时熵产生速率受三方面因素的影响：传热、传热速率和热量与水分传递间的动力学耦合效应，那么为了提高皮革的湿热稳定性，减少皮革的热收缩率就要减少热量产生的速率，也就是说，在动力学方面提高皮革湿热稳定性的关键在于减少传热和传热速率，而且应该尽可能的消除动力学耦合效应^[36]。因此周鲁从动力学角度指出提高皮革的防水性，可以增加皮革的热收缩时间，提高皮革瞬时承受高温的能力。Covington 通过比较未经鞣制的皮粉，经铬鞣制的皮粉和铝鞣制的皮粉的收缩温度，分析认为胶原在受热时状态的变化与胶原分子间的氢键有关，他认为皮粉经鞣制后胶原分子的结构和分布发生了改变，鞣制过程中鞣剂

与胶原分子形成的交联破坏了胶原分子原有的交联,改变了胶原纤维的分散状态,在热作用下,胶原分子链上的氢键作用会减弱,如果温度过高的话氢键就会发生断裂,因此经鞣制的皮粉,胶原分子链上的氢键数目减少,从而胶原蛋白对热的敏感变小,导致皮粉湿热收缩的温度提高^[37]。

国内一些学者研究了皮革的干热稳定性和湿热稳定性。

皮革在热和湿的共同作用下会发生物理机械性能的变化,例如皮革制品在穿着过程中,因为人体发出的热量产生的水汽会使皮革内的游离酸水解而损伤皮革。因此皮革的湿热稳定是判定皮革耐用性能的一个重要指标。各国对皮革的湿热收缩温度都作了相应的规定,我国规定皮革的湿热收缩温度不得低于95℃^[38]。关于皮革湿热稳定性的研究,有研究者指出了皮革的湿热稳定性与皮革的湿含量有关,在一定范围内皮革的湿热稳定性随着湿含量的增加而出现下降;也有一些研究了皮革制作工序中,如复鞣、加脂等工艺对皮革湿热稳定性的影响。通过收缩温度测定仪和差示扫描量热仪等仪器测试了加入不同的复鞣剂或者加脂剂对皮革湿热稳定性的影响。

但皮革不仅受到湿热作用,在很多情况下也受到干热因素的影响,例如制鞋过程中的热定型,或者夏天穿皮鞋在高温的柏油路上行走等。刘婕,汤克勇等人用热台偏光显微镜研究不同皮粉在干热条件下收缩性能。通过比较未鞣制的皮粉和铬鞣制皮粉的干热收缩曲线得出两个结论:(1)铬鞣制皮粉的开始收缩温度比未鞣制的低,最终的收缩温度比较高;(2)未鞣皮粉最终收缩率比鞣制皮粉高,高达45%,并且铬含量越高这两种现象表现的越明显。在铬离子浓度相同时不同种类的铬鞣剂对皮粉干热稳定性的影响不同,糖还原鞣液的鞣制效果比氯化铬鞣液的要好;同时研究了皮革生产工艺过程中的一些工序对皮革热稳定性的影响:首先比较了干燥和未干燥皮粉的收缩量发现,经过干燥皮粉的收缩量在220℃以前大于未经干燥的皮粉,但其最终收缩率小于经过干燥的皮粉;然后研究了鞣制工艺中采用不同的鞣制剂,根据鞣剂渗透的均匀度及与胶原结合的牢固度为评判标准,甲醛、戊二醛的渗透比较好,而葡萄糖还原铬鞣液的结合最为稳定;接着通过实验对比发现不管采用什么复鞣剂,复鞣后皮粉的热稳定进一步提高了;加脂工序中无论采用什么类型的加脂剂,皮革的干热收缩温度都下降了,但收缩率也有所下降,可能是加脂剂的加入虽然使胶原蛋白分子间的距离增大,减少了分子间的作用力,但是同时也减少了胶原蛋白可以回缩的体积,所以总的来说加脂工艺对皮革热稳定性是有所提高的^{[39],[24]}。

1.3 本课题的研究内容及意义

从前面的介绍我们知道,无论是在皮革的加工、生产还是使用过程中,皮革的热稳定性都是一项很重要的性能,如何提高天然皮革的热稳定性也受到越来越多的关注,许

多研究者都从事着皮革热稳定性的研究工作。然而目前关于这一方面的研究是以传统的实验方法为主要的研究手段，如差示扫描量热法（DSC）、差热分析法（DTA）、热重分析法（TG）、差示扫描量热法（DTG）、热台偏光显微镜法、质谱法、X射线衍射法等，他们从不同的角度探讨了皮革的结构和热稳定性之间的关系，并通过大量的对比试验研究了皮革热稳定性的各种影响因素，如水分，鞣制，复鞣，加脂等。但是传统的方法存在很多弊端，例如，要得到实验结果必须进行大量的对比试验，因此需要花费很多时间，耗费很多实验材料；或者有的实验条件目前可能无法达到而限制了实验的进行，比如一些极端的温度或者压力；再者皮革实验所用到的一些材料是会对环境造成污染的，例如铬等鞣制剂，这样就不利于环保。

本文试图从微观结构入手，采用分子动力学模拟的方法，通过 TINKER 软件对 1cgd 这种 I 型胶原蛋白在 Amber 99 力场，自由边界条件和正则系综（NVT）的条件下，分别在 260K，300K，340K 和 380K 的温度下进行 1 纳秒的分子动力学模拟，时间步长为 1 飞秒，建立 1cgd I 型胶原蛋白在不同温度下的分子模型。然后通过均方根偏差、回转半径和氢键这三个参数以及不同时刻的轨迹截取图，对模拟得到的模型轨迹进行分析，研究 1cgd I 型胶原蛋白的空间构象随着温度的变化而发生的改变，从微观层面研究天然皮革的热稳定性，试图为天然皮革热稳定性的研究提供新的理论研究基础。

2 分子动力学

自从 1957 年 Alder 等人在硬球模型下,首次使用分子动力学的方法研究气体和液体的状态方程并取得尝试性的成功以后,就创造了研究物质宏观性质的一种新方法—分子动力学模拟方法^[40]。早期由于受到计算机技术的限制,如速度太慢和内存太小等原因,使得早期的模拟在时间和空间上受到限制。直到 20 世纪 80 年代后期,计算机技术的快速发展和多体式函数的出现为分子动力学模拟技术的进一步发展创造了条件。1972 年 Lee, Edwards 等人发展了分子动力学模拟方法并扩展了其应用范围,使其应用到存在速度梯度的非平衡态系统^[41]。随后相继完成了恒压分子动力学、第一性原理分子动力学、巨正则系统分子动力学等的创建^[42]。

分子动力学模拟,也就是 molecular dynamics simulation,即 MDS,是利用理论方法和计算机技术相结合,对原子核和电子所构成的多体体系中的微观粒子之间的相互作用和运动进行模拟,在这个期间假设每一个原子核都是在全部其他的原子核和电子所构成的经验势场的作用下按照牛顿定律进行运动,这样就能得到体系中粒子的运动轨迹,再按照统计物理的方法计算得出物质的宏观性能^{[43][51]}。它是一种在微观领域模拟原子和分子,进而评估和预测材料结构或者性质的一种重要模拟方法,简单的来说就是应用力场,以及根据牛顿运动力学原理所发展的一种计算机模拟方法^{[43][51]}。分子动力学模拟的目的是希望通过计算机模拟了解结构特性和微观相互作用之间的关系,它是常规实验的一种补充方法,可以让我们掌握用其他一些方法所不能得到的新东西。分子动力学模拟的作用有两个:一个是作为微观长度、时间尺寸和宏观世界的桥梁;另一个是作为理论和实验之间的桥梁。

2.1 基本原理

在分子动力学模拟中原子的运动轨迹是通过解牛顿第二运动定律的微分方程得到的,方程如下:

$$\vec{a}_i = \frac{\vec{F}_i}{m_i} \quad \text{式 (2-1)}$$

将牛顿运动方程式对时间进行积分可以得到第 i 个原子 t 时间后的速度和位置:

$$\vec{a}_i = \frac{d^2\vec{r}_i}{dt^2} = \frac{d\vec{v}_i}{dt} \quad \text{式 (2-2)}$$

$$\vec{v}_i = \vec{v}_i^0 + \vec{a}_i \quad \text{式 (2-3)}$$

$$\vec{r}_i = \vec{r}_i^0 + \vec{v}_i^0 t + \frac{1}{2} \vec{a}_i t^2 \quad \text{式 (2-4)}$$

式中: m_i - 第 i 个原子的质量;

a_i - 第 i 个原子在 t 时刻的加速度;

F_i - 第 i 个原子在 t 时刻所受的力;

v_i - 第 i 个原子在 t 时刻的速度矢量;

r_i - 第 i 个原子在 t 时刻的位置矢量;

0 - 各物理量的初始值

所得结果是一条运动轨迹, 它表明了原子的位置和速度随时间发生变化的规律。其中原子所受的力 F_i 可以通过势能函数 U 的梯度求出:

$$\vec{F}_i = -\nabla_i U = -\left(\vec{i} \frac{\partial}{\partial x_i} + \vec{j} \frac{\partial}{\partial y_i} + \vec{k} \frac{\partial}{\partial z_i}\right) U \quad \text{式 (2-5)}$$

从上面两个方程式我们可以得出, 只要给出某原子的初始速度和位置, 然后对该原子在很小时间间隔内作数值积分, 就可以得到该原子任意以前或以后时刻的速度和位移, 也就是得到了原子的运动轨迹(trajectory)。

2.2 运动方程的求解

为了得到原子的运动轨迹, 可以采用各种有限差分法来求解运动方程, 常用的有以下几种算法:

Verlet算法是在60年代后期出现的, 目前使用最广泛的利用数值积分求解运动方程组的算法, 它的优点是编程比较容易并且占计算机的内存比较小, 缺点是: (1) 准确度不高, 这是因为位置是较小的项与非常大的两项之差得到的, 因此容易造成大的误差; (2) 该算法中必须通过下一步的位置才能计算速度, 对有些需要知道速度的实验就不太方便; (3) 求解下个时刻的位置这一项, 必须知道此刻和前一刻的位置才能得到^[44]; Swope提出的Velocity - Verlet算法目前应用比较广泛, 它可以同时给出位置、速度和加速度, 相对于Verlet算法而言这种算法的优点是准确度高, 可以直接显示速度项, 并且计算量适中^[45]; Hockney提出的Leap - frog “蛙跳” 算法是Verlet算法的一种变化, 这种算法与Verlet算法相比有两个优点: (1) 直接显示速度项; (2) 收敛速度快, 计算量小。但是这种算法也存在明显的缺陷: 位置和速度不能同时显示^[46]; Beeman提出的Beeman算法也是一种与Verlet算法相关的一种算法, 它运用了比Verlet算法复杂得多、更精确的速度表达式, 能更精确的计算出动能, 由于它的表达式比较复杂所以相对于Verlet算法, 计算量较大^[47]; Gear提出的基于预测-校正积分方法的Gear算法分三步来完成: (1) 根据Taylor公式展开, 预测原子新的位置、速度和加速度; (2) 根据计算出的力来计算加速

度; (3) 得出的加速度与Taylor展开式中的加速度进行比较, 在校正步里用两者之差来校正位置和速度项, 这种算法的缺点是占有计算机的内存较大^[48]; Rahman算法实际上是Gear算法的一种变形, 虽然可以得到比较精确的结果但是计算量太大而很少使用^[49]。

不管是哪种算法都是按照以下步骤进行的: (1) 设定粒子的初始速度和位置; (2) 将时间分成很多份, 每一份设定为 Δt ; (3) 计算粒子在 $(t+\Delta t)$ 时刻的受力; (4) 根据受力情况计算粒子的加速度; (5) 由加速度计算出粒子在 $(t+2\Delta t)$ 时刻的速度和位置; (6) 从第三步开始循环计算粒子下个时刻的速度和位置, 直到得到所要模拟的所有状态。

本文为了得到更加精确的结果, 选取的是Beeman算法, 将在下一章中介绍。

2.3 势函数

从分子动力学的定义我们知道进行分子动力学模拟要计算微观粒子间的相互作用, 因此选择适合的势函数是进行分子动力学模拟的基础, 势函数选取的越恰当粒子间的相互作用就计算的越精确, 那么模拟得到的体系就更加接近于实际的体系。

原子间的势函数是从对势向多体势的发展: 对势包括 1957 年 Alder 和 Vainwright 在分子动力学模拟中采用的间断对势^[50]和 Rahman 在 1964 年对氢元素的研究中应用的非间断的对势, 它认为原子间的相互作用只是一个原子与另外一个原子两两之间的相互作用, 而与其他原子的位置无关^[51]; 但是在一个存在多原子的体系中, 一个原子的位置不同将影响其它原子间的相互作用, 所以 20 世纪 80 年代 Daw 和 Baskes 首次提出 EAM 势^[52], 随后很多研究者在 EAM 势的基础上相继提出了其他的一些多体势函数等, 这些多体势能够更加准确地表示多原子体系^[53]⁶⁷。

本研究选取的是 AMBER 力场下的经典对势, Lennard-Jones 势, 下一章中会介绍。

2.4 一些重要术语

(1) 系综

系综的概念是由约西亚·威拉德·吉布斯 (J. Willard Gibbs) 在1878年提出的, 它是指在一定的宏观条件下, 大量处于各种运动状态并且各自独立, 但是性质和结构完全相同的系统的集合。平衡态分子动力学模拟总是在一定的系综下进行的, 常用的平衡系综包括微正则系综, 正则系综和等温等压系综^[53]⁶⁸。

在平衡系综中, 把能量、体积和粒子数都固定的系统称为微正则系综, 也称为NVE系综, 它是保守、孤立的系统系综, 因为在分子动力学模拟的过程中, 体系与外界不发生能量交换, 粒子数也守恒, 体系中的体积也不发生变化, 也就是说系统中的粒子数N, 体积V和能量E都保持不变; 正则系综则是在分子动力学模拟的过程中保持系统中的粒

子数 N ，体积 V 和温度 T 不变，并且总动量为零，所以也叫做NVT系综；等温等压系综也就是NPT系综，它是在分子动力学模拟的过程中保持系统中的粒子数 N ，压力 P ，温度 T 不变的系综^[54]。

平衡系综的控制方法有很多，比如在正则系综（NVT系综）和等温等压系综（NPT）中要保持系统的温度不变，可以通过直接对速度进行标度来保持系统的温度不变，这种方法可以使系统很快的达到平衡，或者是采用Berendsen热浴法，也就是让系统和恒温槽通过热交换来保持系统的温度不变，这种方法很简单，应用的时候也很方便^[55]。在等温等压系综（NPT）中要保持系统的压力不变，一般通过改变系统的体积来实现系统的压力不变，如采用Berendsen压浴法和Parrinello-Rahman（P-R）^[56]方法。本实验选择的是正则系综。

（2）边界条件

分子动力学是对大量粒子运动的一个统计过程，但计算机只能处理有限的粒子，因此为了减少计算量，会从庞大的体系中选取某个区域的粒子来代表整个体系进行模拟，在这个时候就要考虑所截取的这个区域中处于边界位置的粒子，因为它们还会受到外界的作用力，所以分子动力学模拟中一般会使用边界条件。一般使用的周期性边界条件有两个优势：一是保证系统粒子数的恒定，也就是粒子在运动的时候，当有粒子运动到所设定的模型以外时，则通过设定条件，使有相同数目的粒子回到所设定的模型中，这样就保证了所模拟系统粒子数目的恒定，即密度保持不变；二是在计算原子间作用力的时候通过采取最近镜像方法，采用截断半径 r_c (cut-off radius)，也就是限定一个粒子只与该粒子所处的基本单元中的其他粒子或者与它最相邻的粒子间发生相互作用，这样做的目的是消除边界效应，保证各原子全面受力^[57]⁴⁴。根据研究的需要还可以选择固定性边界条件和自由边界条件，本研究选择的是自由边界条件，也就是对边界上的粒子不限制任何条件，这样可以使模拟更加准确，但是计算量会加大。

（3）积分步长

积分步长的选取在分子动力学的模拟中是一个十分关键的步骤，它的选取决定了模拟的准确性和模拟时间的长短。积分步长选取越大计算速度也就越快但是会导致较大的截断误差，甚至会出现数值解的不稳定，降低计算的准确性；积分步长选取越小，计算的准确性越高但是却增加了计算次数，模拟的时间也就会越长。因此，为了合理的选取积分步长，一般取系统中最快运动粒子运动周期的十分之一作为模拟体系的积分步长值，因为在这个时间段内才能保持体系的速度和加速度的恒定^[57]⁵²。

2.5 分子动力学的应用

分子动力学模拟由于它独特的优势在许多领域中得到广泛地应用，如物理、化学、

生物和材料科学等，这里仅介绍分子动力学在研究蛋白质稳定性方面的应用。

曹同成等人用分子动力学的方法模拟了微波加热升温过程对水溶液中蛋白质构象的影响，结果表明血红蛋白在27℃加热到40℃时蛋白质结构已经开始转化为松散式，随着温度的进一步升高这种转化就越来越明显，说明血红蛋白的热稳定性不太好，而钙调蛋白在温度为70℃以下时结构没有明显的变化，证明钙调蛋白的热稳定性比较好^[58]；蒲尚志采用分子动力学方法用牛皮胶原的氨基酸序列置换胶原模型多肽PPG，得到牛皮胶原模型多肽BOV，研究了不同温度对BOV热稳定性的影响，结果显示：（1）在相同温度下BOV的根均方涨落都高于PPG，随着温度的升高，根均方涨落的变化小于PPG，说明氨基酸序列对蛋白质的柔韧性有影响，BOV的柔韧性比PPG高；（2）在相同温度下BOV的根均方偏离都高于PPG，随着温度的升高，根均方偏离的变化小于PPG，同时PPG和BOV的氢键数目随着温度的升高都减少了，但是BOV减少的更多，说明氨基酸序列对蛋白质的热稳定性有影响，脯氨酸含量低的BOV的热稳定性比脯氨酸含量高的PPG的热稳定性差^[59]；Clark用分子动力学方法研究了尿素对水溶液中S肽链蛋白质结构的影响，通过比较氢键数量发现当尿素浓度较低时主要是S肽链与水分子间形成的氢键和S肽链内的氢键维持蛋白质的稳定性，S肽链与尿素之间形成的氢键数目比较少，而且氢键只是在S肽链的侧链上形成的，起到稳定肽链的作用，因此S肽链蛋白质在尿素浓度较低时稳定性有所提高，随着尿素浓度的升高，尿素不仅与S肽链侧链上的极性氨基酸形成氢键，还会渗透到S肽链的内部与骨架上的氨基酸残基形成氢键，这样会破坏S肽链骨架氢键而使S肽链变性^[60]；一些研究者用分子动力学方法研究了不同渗透剂对胰凝乳蛋白酶抑制剂2稳定性的影响，结果表明，不管渗透剂是什么类型都会提高胰凝乳蛋白酶抑制剂2的稳定性，只是提高的程度有所不同，渗透剂的极性表面积分率与渗透剂的稳定性相关，当极性表面积分率大于0.7时呈正相关，小于0.7时负相关；渗透剂的体积则与胰凝乳蛋白酶抑制剂2的稳定性成正比^[61]。

2.6 本章小结

从分子动力学的介绍我们了解到分子动力学是通过模拟微观粒子之间的相互作用和运动得到粒子的运动轨迹，通过统计力学对粒子的运动轨迹进行分析，进而统计出物质的宏观性能，也就是说分子动力学方法是联系宏观和微观的重要工具。而目前也有很多研究者运用分子动力学方法研究蛋白质的各种性质，如前面介绍的蛋白质稳定性，因此，用分子动力学模拟的方法研究天然皮革最基本的物质-胶原蛋白的热稳定性是可行的，这也是本文的一个创新点。我们通过分子动力学模拟不同温度下胶原蛋白微观粒子间的相互作用和运动，从而得到不同温度下胶原蛋白的运动轨迹，然后通过统计力学的分析就可以从皮革原子水平的空间构象来研究皮革的热稳定性。

3 实验部分

3.1 实验所用软件

本实验所用的软件包括 TINKER 软件, Force Field Explorer 软件和 MATLAB 软件, 前两个软件用于对 1cgd I 型胶原蛋白进行分子动力学模拟得到原子的轨迹文件, 后一个软件用于绘图和进行数据分析。同时还用到了 Compaq Visual.Fortran.Version.6.6 编译器对模拟轨迹进行分析。

3.1.1 TINKER 软件

本实验所采用的模拟软件为 TINKER 软件(<http://dasher.wustl.edu/tinker/>)。

TINKER 是 Jay William Ponder 教授开发的一个比较完整的分子模拟软件包, 它里面包含了很多程序, 而不是指某一个单独的程序。TINKER 软件是一个用于算法开发和参数化设置的平台, 可以用于分子力学和分子动力学的计算, 特别适用于计算生物聚合物, 它可以在源代码的基础上根据实验的需要自己写一些程序或者对程序进行修改。TINKER 可以使用许多力场及参数, 目前可以使用的力场有 AMBER ff94/ff96/ff98/ff99, CHARMM 19/22/27, Allinger MM2-1991, Allinger MM3-2000 和 OPLS-AA 等, 其它的力场可以通过添加一些新的参数文件而在 TINKER 中运用。

TINKER 的功能主要包括: (1) 通过共轭梯度法、可变动的度量标准或者简洁化的牛顿方法, 采取卡迪尔直角坐标, 扭转角或者刚体空间, 进行能量最小化; (2) 通过限定周期性边界条件, 控制温度和压力, 进行随机分子动力学; (3) 采用 RESPA 集成复杂的时间步长用于更加高效的分子动力学模拟; (4) 正常模式下的振动分析; (5) 包含高效随机成对度量化的几何距离法; (6) 建立蛋白质和核酸的序列结构; (7) 单点势能的分析和分解; (8) 核查和分析标准的和用户自定义的势能; (9) 定位两个极小态之间过渡状态的位置; (10) 通过构象扫描方法进行能量表面搜索; (11) 通过自由能微扰工具或者加权直方图分析计算自由能; (12) 分析和比较静电势能; (13) 使分子间的势能与结构和热力学的参数相吻合; (14) 通过卡罗蒙特法和模拟退火法进行全局优化, 通过势能平滑和搜索 PSS 方法使能量表面平滑。本研究主要是使用 TINKER 软件进行能量最小化和分子动力学模拟。

TINKER 软件的操作系统有很多种, 包括 Windows, PC Linux, SGI, Macintosh, RS/6000, Alpha, HP-UX, Sun 等。本实验是在 Windows 操作系统下采用 TINKER 5.1

软件进行分子动力学模拟的。

3.1.2 Force Field Explorer 软件

Force Field Explorer 软件(<http://dasher.wustl.edu/ffe/>), 简称 FFE, 是与 TINKER 软件相关的一款软件, 是 TINKER 软件一个最基本的辅助包, 它是为 TINKER 设计的一个完整的 JAVA 图像用户界面, 为了更方便的使用 TINKER 中的某些功能。该软件设计的目的是为计算化学和分子工程提供一个更加稳定的环境, 并不是简单的把它作为一个三维结构图形的查看器。目前的功能包括: 安装和实时可视化建模程序, 包括全局优化或者部分优化, 抽样方法, 叠加等; 基本的三维分子可视化, 一般的显示方式包括线框, 球形, 棒形, 管状等; 可以设定一些力场的特殊参数, 例如 Lennard-Jones 半径以及部分电荷的数量; 拥有一个可以控制建模参数的算法, 功能强大但操作简单的关键词的编译文件; 可以对 TINKER 软件所产生的轨迹文件进行观看和重放。

从图 3.1 Force Field Explorer 软件的界面截图我们简单地了解一下该软件。

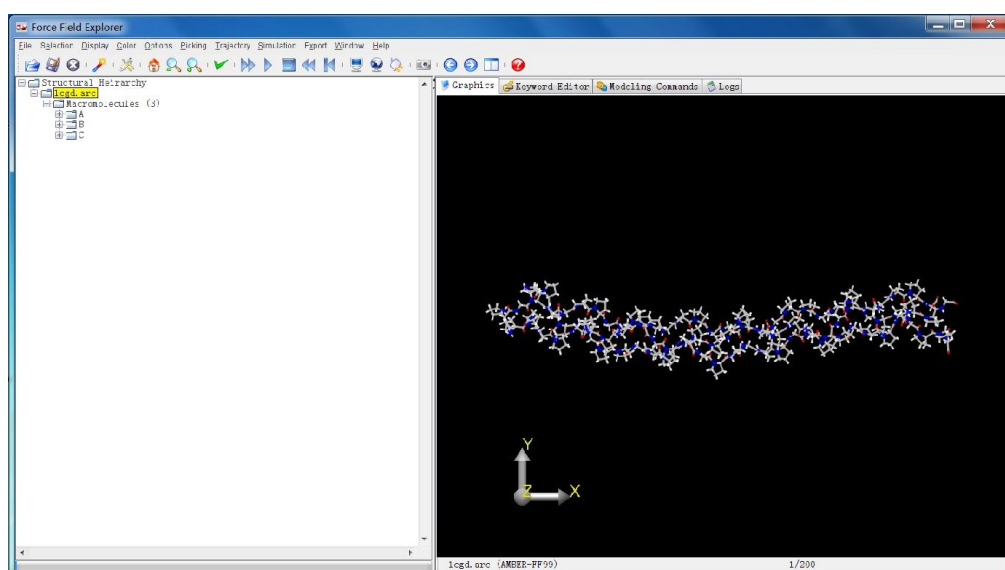


图 3.1 Force Field Explorer 软件截图

Fig 3.1 A screenshot of Force Field Explorer

该软件整体上包含三大部分: 第一部分是截图上面两行, 也就是标注 File、Selection、Display 等的这一行和它的下一行, 这一部分就是该软件的菜单部分, 也就是通过该软件可以执行的操作, 第二行是一些常用操作的按钮。本实验主要用到的是 Trajectory 这个操作, 也就是可以观察蛋白质的空间构象并对蛋白质构象的轨迹进行截取, 这个功能在第二行同样有可操作的按钮; 第二部分是截图左方的 Structural Hierarchy 部分, 也就是蛋白质分子的结构层次, 一般形象的称为分子树, 用于导航和选择。这一部分能详细

的了解你所选择的蛋白质分子的结构,如图 3.1 中就可以知道所选择的 1cgd 这个蛋白质有 A、B、C 三条链,当然点击这三个字母前面的“+”号就可以进一步了解它的残基数目等更加详细的信息;第三部分是截图右方的 Graphics 部分,也就是视图部分,从这一部分我们可以观察到分子动力学模拟的轨迹。试图左下角用黄色标注的 X/Y/Z 为视图的方向轴,也称为全球轴,当左键点击全球轴后可以拖动整个图像围绕原点进行转动,我们也可以通过按住鼠标左键拖动 X/Y/Z 方向轴来从不同角度观察蛋白质分子。Graphics 后面的 Keyword Editor 可以用来修改控制 TINKER 软件各方面计算的关键词。Modeling Commands 是力场建模指令,可以执行 TINKER 软件大部分的程序。Logs 为记录面板,它是一个简单的文本编译器。X/Y/Z 方向轴下方左边的 1cgd.arc (AMBER-FF99) 表明我们是在 AMBER 99 力场下观察的 1cgd,右边的 1/200 表示总共有 200 个轨迹,现在显示的为第一个,从这里我们知道 Force Field Explorer 软件可以根据研究的需要截取模拟过程中任意时刻的轨迹图。

Force Field Explorer 软件可以打开 XYZ 或者 Archive 文件,在选择了要打开的文件之后,文件并不能马上打开,而是会弹出选择力场的界面,如图 3.2。也就是我们在查看分子动力学模拟结果之前必须先选择你模拟过程中所使用的力场,这就说明力场的选择在分子动力学模拟中的重要性。选择力场后 TINKER 的 Key 文件就会自动创建。



图 3.2 Force Field Explorer 软件力场文件选择的示意图

Fig 3.2 A screenshot of selecting a parameter file in Force Field Explorer

Force Field Explorer 软件的操作系统目前有 Linux 和 Windows,本实验是在 Windows 操作系统下,使用 Force Field Explorer 5.0 软件观察模拟过程中 1cgd I 型胶原蛋白分子运动轨迹的变化,这样分子动力学模拟的结果就可以通过图形来观察 I 型胶原蛋白分子在不同温度下分子构象的变化情况,从而能进一步从微观的层面上理解 I 型胶

原蛋白的热稳定性。

3.1.3 MATLAB 软件

MATLAB 软件 (<http://www.mathworks.com/>) 是美国 Math Works 公司于 1984 年推出的商业数学软件，用于矩阵运算、图形处理、数据分析以及程序设计等，它的名称是由矩阵实验室 Matrix Laboratory 所合成的。MATLAB 自 1984 年推出以来已经从最原始的 1.0 版本升级到了现在的 7.13 版本，总共经历了 34 段升级历程，相应的操作系统也由最开始的 Dos 系统发展为现在的 Windows 7、Vista 等多种操作系统，可见它的发展速度是十分惊人的。MATLAB 之所以发展势头如此迅猛，就在于它具有很多优点，其最大的优点是语言简洁，使用方便。MATLAB 语言是基于 C++ 语言基础，而且比 C++ 语言更简洁，即使是非计算机专业的研究人员也可以方便使用。这种语言还具有调试性好和扩充能力强的特点，也就是说在 MATLAB 中编辑的程序可以在其他操作系统中运行。除此之外，它还有很多优势，如强大的数据处理能力，出色的图形处理能力等。本实验采用 MATLAB R2011b 编写一些程序用于图形的绘制和数据的统计。图 3.3 为 MATLAB R2011b 软件界面的截图。

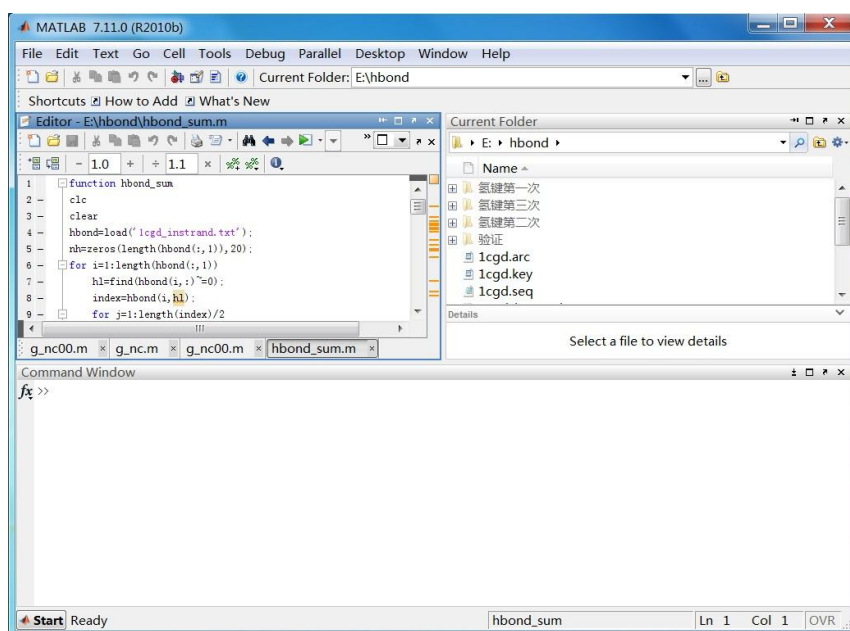


图 3.3 MATLAB R2011b 软件界面截图

Fig 3.3 A screenshot of MATLAB R2011b

MATLAB R2011b 软件中我们使用最多的是图 3.3 下方的“Command Window”窗口，也就是命令窗口，这一部分是主要程序输入和运行的地方，它可以显示除了图形以外的所有运算结果，本研究中氢键的数据统计结果部分就是在命令窗口部分显示的。实

验中图形的绘制以及数据的统计部分是在图 3.3 左上方的编辑器里面进行的程序命令的编辑。

3.2 实验前期准备

3.2.1 PDB 文件的准备

PDB 文件需要从蛋白质数据库(<http://www.rcsb.org/pdb/home/home.do>), 即 Protein Data Bank (PDB) 中免费下载。

蛋白质数据库是 1971 年美国布鲁海汶 (Brookhaven) 国家实验室的 Walter Hamilton 建立的, 是一个专门收录蛋白质及核酸三维结构资料的免费数据库。在 1998 年 10 月, PDB 被移交给了结构生物学合作研究协会 (Research Collaboratory for Structural Bioinformatics), 也就是我们现在用到的 RCSB PDB。最初蛋白质数据库中仅有 7 个蛋白质结构, 但它的增长速度是非常惊人的, 截至 2008 年 2 月蛋白质数据库中已有接近 50,000 个原子分辨率的蛋白质及相关复合物的三维结构坐标。

从蛋白质数据库中下载之前必须知道你所需要下载的蛋白质 ID, 也就是说, 每一种蛋白质的三维结构在该数据库中都有一个特定的编号, 就像人的身份证一样, 因此我们首先要找到天然皮革在蛋白质数据库中的编号。在绪论中已经指出了目前还很少有人做天然皮革的分子动力学模拟, 因此相关的资料很少, 查阅大量资料后也无法确定每一种动物皮革的 ID 号, 因此决定研究天然皮革的大类。通过查找文献最终确定从蛋白质数据库中下载蛋白质编号为 1CGD 的蛋白质 pdb 文件。

本实验选择编号为 1CGD 的蛋白质作为研究对象主要有以下两个方面的原因: (1) 1CGD 是一种 I 型胶原蛋白, 前面已经介绍了成年动物皮中含量最高的就是胶原蛋白, 尤其是 I 型胶原蛋白^[62], 在猪皮中的含量达到 80%-85%^[63], 在牛皮中的含量也很高, 所以选取 1CGD 作为研究对象具有代表性; (2) 考虑到本实验所用到的电脑配置, 太复杂的蛋白质计算机运行的太慢, 而 1CGD 相对而言比较简单, 所以选择编号为 1CGD 的这种蛋白质作为研究对象比较合理。

J.BELLA 等人通过 X 射线衍射法在分辨率为 1.85 Å 时得到 1CGD 的空间构象示意图, 如图 3.4 所示。1CGD 的二级结构主要包括三个 α 螺旋和螺旋外的一些无规则卷曲。它的分子量为 8535 道尔顿, 包含 A、B、C 三条 L 型的氨基酸肽链, 蛋白质分子的序列总长度为 90 个残基, 每条链上都有 30 个残基, 而且三条链的氨基酸序列都是 PRO HYP GLY PRO HYP GLY PRO HYP GLY PRO HYP GLY PRO HYP ALA PRO HYP GLY PRO HYP GLY PRO HYP GLY PRO HYP GLY PRO HYP GLY。其中 PRO 代表脯氨酸, 它是常见氨基酸中唯一的环状亚氨基酸; HYP 代表羟脯氨酸, 是脯氨酸羟化后的产物,

也是一种亚氨基酸；ALA 代表丙氨酸，一种脂肪族的非极性氨基酸。

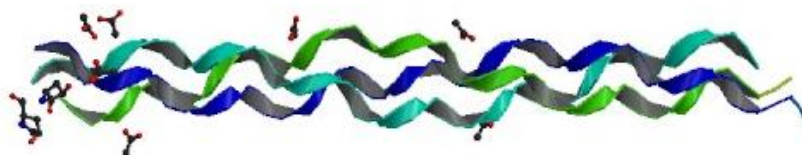


图 3.4 1CGD 蛋白质天然态的构象示意图

Fig 3.4 The native conformation of 1CGD

通过观察 1CGD.pdb 文件我们发现该文件中有些对模拟的实验结果不会产生很大的影响，但是本身不能被 TINKER 软件所识别，而且考虑到后面的模拟条件，在专业技术人员的指导下我们对 pdb 文件进行前期处理：第一点，该文件中包含两种原子基团的坐标，ATOM 和 HETATM，ATOM 定义的是标准原子基团的坐标，而 HETATM 是非标准的，TINKER 软件不能识别，因此需要把 HETATM 更换为 ATOM；第二点，考虑到模拟过程中采用的是 GB/SA 类型的溶剂模型，也就是隐含水模型，因此可以对 1CGD.pdb 文件进行去水处理以减少计算量；第三点，去掉 1CGD 的开端或者尾端的 ACY 帽子氨基酸。虽然这些处理会使 1cgd 减少一些氨基酸，但是通过对比试验我们证明了这些处理不会对蛋白质的性质造成较大的影响，也不会对文中所研究的内容和结果分析产生很大的偏差，只是减少了计算量，缩短了模拟的时间。

综上，pdb 文件的准备过程有四个步骤：（1）从蛋白质数据库中下载 ID 为 1CGD 的蛋白质结构；（2）用 ATOM 替换 HETATM 类型的原子基团坐标；（3）去除 1CGD 中的水分子；（4）把 1CGD 中的帽子氨基酸 ACY 去掉。经过这四个步骤就形成了这次分子动力学模拟的初始结构，为了与下载的 1CGD.pdb 相区别，我们把该结构命名为 1cgd.pdb。

3.2.2 Key 文件的准备

key 文件是分子动力学模拟的控制文件，包括了模拟运行时所需要用到的全部可选参数。key 文件中每行为一个关键词+参数，整个 key 文件对大小写不敏感，需要多个数字做参数时，多个数字可以用空格，逗号或者 Tab 键隔开。本实验 key 文件的设置如下：

Output Control ARCHIVE

Force Field Selection PARAMETERS amber99.prm

SOLVATE GB/SA

第一行的 Output Control 为输出控制, ARCHIVE 为 TINKER 软件运行后生成的轨迹文件(trajector), 只有在 key 文件中制定 ARCHIVE 关键词才能生成。则第一行表示运行 TINKER 要输出原子的轨迹文件。

第二行的 Force Field Selection 为力场的选择, PARAMETERS 为参数的设置, prm 表示文件为参数文件, TINKER 力场参数文件都是以 prm 为扩展名的。我们选择的是 amber 99 力场。

第三行的关键词 SOLVATE 是提供多种连续溶剂算法, 我们选择的是 GB/SA, 也就是将水环境看成是一个连续介质的隐含水模型。GB/SA 隐含水模型是一种近似模型, 它可以在系统没有水的情况下模拟出有水的效应, 因此可以节省大量的计算时间, 很适合进行分子动力学模拟。

从前面的介绍我们知道选择适合的势函数是进行分子动力学模拟的基础, 而势函数又是由分子力场的类型所决定的, 因此分子力场的选择非常重要, 分子力场是进行分子动力学计算的基础, 这一点从前面介绍 Force Field Explorer 软件时也说明了。一般分子力场包含两类相互作用: 化学键相互作用和非键相互作用。化学键相互作用也就是分子内相互作用, 主要包括: (1) 键伸缩能(Bond stretching): 分子的所有化学键在键轴方向上的伸缩运动产生的能量变化; (2) 键角弯曲能(Angle bending): 由分子键角的改变产生的能量变化; (3) 二面角扭曲能(Torsion rotation): 单键旋转引起分子骨架扭曲所产生的能量变化。非键相互作用项也就是分子间的相互作用, 主要是指静电力作用(Electrostatic)和范德华相互作用(Van der Waals)。

本实验是选择 AMBER 力场进行分子动力学模拟的。AMBER 力场是 Weiner 等人 1984 年发展的在生物大分子的模拟计算领域有着广泛应用的一个分子力场, 它通过拟合与蛋白质和核酸体系具有类似结构单元的有机小分子实验数据来拟合其相应的力场参数。AMBER 力场是一种全原子力场, 也就是它精确定义了体系内的每一个原子的参数, 并且该力场的参数大多数来源于实验结果。公式 3-1 是 AMBER 力场势函数的形式。

$$V = \sum_{\text{bonds}} \frac{k}{2} (l - l_0)^2 + \sum_{\text{angles}} \frac{k}{2} (\theta - \theta_0)^2 + \sum_{\text{torsions}} \frac{V_n}{2} [1 + \cos(n\omega - \gamma)] \\ + \sum_{i < j} \left\{ 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right\} \quad \text{式 (3-1)}$$

在 AMBER 力场中, 前三项为化学键相互作用, 第一项和第二项是用谐振子模型描述的键伸缩能和键角弯曲能; 第三项是用傅里叶级数形式描述的二面角扭曲能; 第四项为非键相互作用, 分别为用 Lennard-Jones 公式描述范德华相互作用和用库仑公式描述

的静电力作用^[64]。

我们选择的 amber 99 力场包括了以下的几个部分：

(1) Atom Type Definitions 也就是原子类型的定义。因为一个蛋白质分子中包含很多不同种类和数目的原子，不可能对所有的原子类型形成的键弯曲，二面角等参数进行参数化，所以就把原子分成几种不同的原子类别(Atom Class)，原子类别下面又定义了不同的原子类型(Atom Type)。这一部分把 TINKER 中的原子类别转化为 Amber 定义下的原子类型，这样属于同一原子类别下的不太重要的原子都使用同一参数，这样就减少了需要参数化的原子类型的数目，图 3.5 显示了 amber 99 力场对原子的定义；

```
#####
##  TINKER Atom Class Number to Amber Atom Types  ##
##                                                    ##
##   1  CT      11  CN      21  OW      31  H0      ##
##   2  C       12  CK      22  OH      32  HS      ##
##   3  CA      13  CQ      23  OS      33  HA      ##
##   4  CM      14  N       24  O       34  HC      ##
##   5  CC      15  NA      25  O2      35  H1      ##
##   6  CV      16  NB      26  S       36  H2      ##
##   7  CW      17  NC      27  SH      37  H3      ##
##   8  CR      18  N*      28  P       38  HP      ##
##   9  CB      19  N2      29  H       39  H4      ##
##  10  C*      20  N3      30  HW      40  H5      ##
#####
```

图 3.5 amber 99 力场对原子的定义

Fig 3.5 The definition of atom type in amber 99

- (2) Van der Waals Parameters 范德华参数；
- (3) Bond Stretching Parameters 键的伸缩参数；
- (4) Angle Bending Parameters 角的弯曲参数；
- (5) Torsional Parameters 二面角参数；
- (6) Atomic Partial Charge Parameters 原子部分电荷参数；
- (7) Biopolymer Atom Type Conversions 生物高分子原子类型的转化。

3.3 实验过程

我们在 DOS 命令下运行 TINKER 软件：

- (1) C:\User\sun> E:
- (2) E:\> cd tinker
- (3) E:\tinker> pdbxyz.exe
- (4) Enter Protein Data Bank File Name: 1cgd.pdb
- (5) Enter the Chain Names to Include(A B C [ALL]): ALL
- (6) E:\tinker> minimize.exe
- (7) Enter Cartesian Coordinate File Name: 1cgd.xyz
- (8) Enter RMS Gradient per Atom Criterion: 0.01
- (9) E:\tinker> dynamic.exe
- (10) Enter Cartesian Coordinate File Name: 1cgd.xyz_2
- (11) Enter the Number of Dynamics Steps to be Taken: 1000000
- (12) Enter the Time Step Length in Femtoseconds: 1.0
- (13) Enter Time between Dumps in Picoseconds: 5
- (14) Available Simulation control Modes:
 - 1) Constant Total Energy Value (E)
 - 2) Constant Temperature Via Value
 Enter the Number of the Desired Choice: 2)
- (15) Enter the Desired Temperature in Degrees K: 300

(1) - (2) 步是进入到TINKER所在的文件夹中，在E盘的tinker文件夹中。

(3) - (4) 步是把1cgd.pdb文件转化为TINKER软件所能识别的1cgd.xyz坐标文件。xyz坐标文件是TINKER文件独有的，它在普通的坐标文件基础上添加了原子序号、原子类型和若干列与当前原子联通的原子序号。以1cgd.xyz文件的一部分为例加以解释说明，如表3.1所示。

表 3.1 1cgd.xyz 文件的部分说明

Table 3.1 Section describes of 1cgd.xyz

原子 序号	原子 类型	X 坐标 (Å)	Y 坐标 (Å)	Z 坐标 (Å)	电荷 类型	与当前原子联通的原子序号			
1	N3	3.777000	3.067000	7.824000	404	2	5	6	10
2	CT	4.933000	3.909000	8.279000	405	1	3	7	8
3	C	6.239000	3.357000	7.692000	406	2	4	17	
4	O	6.313000	2.168000	7.380000	408	3			
5	H	4.114999	2.322320	7.231287	407	1			

以上五行为截取的 1cgd.xyz 文件中的前五个原子的 xyz 坐标文件。该文件包含 7 列：第一列为原子序号；第二列为原子类型，这里的原子类型是 amber 99 力场文件中对原子

定义的类型，见图 3.5；第三到五列为原子的 X, Y, Z 坐标，单位为 Å($1\text{Å}=10^{-10}\text{m}$)；第六列为这个原子的电荷编号，也就是电荷类型，它也是由 amber 99 力场文件定义的；后面的几列表示与该原子相连或者有相互作用的原子的序号。例如第一行的 2、5、6、10 表示 1 号原子 N3 与编号为 2、5、6、10 的 5 个原子相连或者有相互作用。

第(4)步后得到了 1cgd.xyz 和 1cgd.seq 两个文件，前者是 TINKER 软件特有的坐标文件，后者是氨基酸序列文件，是在用 pdbxyz.exe 程序将 pdb 文件转化为 xyz 文件时自动生成的。

第(5)步选择所要模拟的链，1cgd 总共有 A, B, C 三条链，我们选择全部进行模拟，因为每条链不是独立存在的，另外两条链会对它产生一定的作用，因此选择模拟所有的链结果比较精确。

第(6) - (8)步是进行能量最小化。能量最小化(Energy Minimization)即寻找研究体系能量极小的状态，是通过对原子坐标的能量函数求极小化得到的，所以也称几何优化(Geometry Optimization)。之所以要进行能量最小化是因为系统内原子之间存在不正常的相互作用，比如系统中有些原子的距离过近时导致局部的能量很高，系统不稳定，可能影响后续的分子动力学过程。消除这种局部高能量的方法就是进行能量最小化。能量最小化实际上是通过改变系统中局部能量过高的原子位置，调整距离过近的原子位置，使得整个系统分布均匀稳定，分子结构处于自然无张力的状态。能量最小化得到的一般是体系的局域能量最小构象，一般把这个构象作为分子动力学模拟的初始构象，因为此时的体系是稳定的。衡量能量最小化的收敛标准是梯度。

能量最小化常用的方法有两种：最速下降法(Steepest Descent Method)和共轭梯度法(Conjugate Gradient Method)，前一种优化方法最简单，它是选取使目标最快速下降的方向，也就是负梯度方向为搜索方向，它的优点是收敛速度快，但是收敛性比较差，一般是用于结构的小修改，用于能量最小化的初始阶段；后一种方法是利用当前的负梯度方向和上一步优化时的负梯度方向来决定下一步的优化，它的特点是收敛的相对效率要高一些，而且需要的能量和计算量也相对于最陡下降法要少一些^[65]。本实验选择共轭梯度法进行能量最小化。

第(7)步是输入笛卡尔坐标，也就是 1cgd.xyz，表示对 1cgd.xyz 进行能量最小化或者说是结构优化。第(8)步是输入 rms 这个收敛标准。rms 是指前后两个原子坐标的均方差(root mean square)，表示坐标的变化。我们对 1cgd.xyz 进行了 3176 步的优化，为了提高收敛的精度我们采用共轭梯度法进行优化，收敛判断的标准是均方根力为 0.01 千卡/(摩尔·埃)(kcal/mol/Å)，也就是均方根力小于 0.01 kcal/mol/Å 时判断能量优化达到了标准，此时能量对坐标变化的梯度接近于 0，认为达到了局部极值，停止优化。能量最小化后生成 1cgd.xyz_2 文件。

第(9) - (15)步是进行动力学模拟。依次输入了笛卡尔坐标:1cgd.xyz_2,也就是我们是对能量优化后的结构进行分子动力学模拟的;模拟的步数(The Number Of Dynamic Steps): 1000000N,模拟步数的选择是很重要的,它决定了整个计算时间的长短,本实验模拟步数的确定是基于对计算机配置的考虑;时间步长(The Time Step Length): 1fs(1fs=10⁻¹⁵s),本文时间步长的选择是依据系统中运动周期最快的氢原子的振动频率;输出频率(Time Between Dumps): 5ps(1ps=10⁻¹²s),这个 5ps 的意思就是每隔 5ps 将体系中各个原子的坐标信息写入文件中;在系综的类型(Available Simulation Control Modes)选择时系统给出了两种选择, Constant Total Energy Value 和 Constant Temperature Via Value,前一种为保持系统的能量一定,后一种为保持系统的温度一定,我们选择第(2)种,也就是我们选择正则系综(NVT 系综);最后我们输入的是实验模拟温度的设定,单位为开尔文(K),因为我们要研究天然皮革的空间结构随温度变化而发生的改变,所以本实验选取了从低温到高温的四个温度条件,分别为 260K、300K、340K 和 380K 进行分子动力学模拟,每 100 步记录一次原子的运动轨迹文件。模拟完后我们得到了 1cgd.arc 文件,arc 文件也是轨迹文件,只不过是 200 个结构的轨迹文件。选择 arc 文件是因为 1cgd 有 1068 个原子,为了方便后面的轨迹分析,我们把 1068 个原子的坐标作为一个整体进行存储,一共存储了 200 个结构的轨迹文件在 arc 文件中。

本实验分子动力学轨迹的得到是基于 Beeman 算法。Beeman 算法是 Verlet 的一种演变,也是由泰勒公式展开得到的,表达式为:

$$\vec{r}(t + \delta t) = \vec{r}(t) + \vec{v}(t) \delta t + \frac{2}{3} \vec{a}(t) \delta t^2 - \frac{1}{6} \vec{a}(t - \delta t) \delta t^2 \quad \text{式 (3-2)}$$

$$\vec{v}(t + \delta t) = \vec{v}(t) + \frac{1}{3} \vec{a}(t) \delta t + \frac{5}{6} \vec{a}(t) \delta t - \frac{1}{6} \vec{a}(t - \delta t) \delta t \quad \text{式 (3-3)}$$

其中 $\vec{r}$ 为位置, $\vec{v}$ 为速度, $\vec{a}$ 为加速度, δt 为时间步长。从 Beeman 算法关于位置和速度的公式可以看出时间步长 δt 是分子动力学模拟过程中非常重要的一个参数, δt 的取值决定了计算的速度, δt 越大计算的速度越快,但当 δt 过大时又会导致积分过程的发散和积分值的不准确。本实验 δt 的取值是选择系统中最高运动频率的氢原子的十分之一作为标准。前面也介绍了 Beeman 算法比较精确,但是计算量很大,所以整个的模拟过程用的时间比较长,每个温度下模拟一次大概需要 2 天的时间。++++*

本实验采取的是自由边界条件,相当于是在真空下进行模拟,因此不需要考虑周期性条件,粒子间也没有采用截断半径的方法来计算远程作用力,系统中的每一个粒子都能看到所有其他的粒子,这种情况下模拟的比较准确。

3.4 模拟轨迹的分析

3.4.1 模拟轨迹的分析方法

前面分子动力学模拟得到的是 1cgd I 型胶原蛋白原子微观水平的坐标，但我们要研究的是 I 型胶原蛋白所表现出来的宏观物理性质，就要用到统计力学。统计力学使用数学方程式表达了宏观状态与系统内部粒子运动之间的关系。

本文研究的是 I 型胶原蛋白的稳定性，我们选取均方根偏差(RMSD)、回转半径(Rg)和氢键(hb)三个参数对 1cgd I 型胶原蛋白在不同温度下结构状态的变化进行分析。

(1) 均方根偏差(Root Mean Square Deviation, RMSD)

均方根偏差反映的是蛋白质分子在进行分子动力学模拟的过程中产生的结构相对于优化后的初始结构的偏离程度，它的计算公式为：

$$\text{RMSD} = \sqrt{\frac{1}{M} \sum_{i=1}^N m_i \|r_i(t_1) - r_i(t_2)\|^2} \quad \text{式 (3-4)}$$

式中： m_i - 第 i 个原子的质量；

M - 所有原子的质量和；

$r_i(t)$ - 第 i 个原子在时间 t 时的位置；

t_1 - 模拟过程中的某一时刻；

t_2 - 一般为初始时刻

计算 RMSD 时可以针对所要分析蛋白质的某一部分，比如蛋白质的全部原子或者骨干部分的原子或者 α 碳原子。本实验选择的是多肽链中主链上 α 碳原子 (C^α) 的均方根偏差作为研究对象，这是因为在前面介绍胶原蛋白时已经知道每个氨基酸残基都含有一个 C^α 原子，对 C^α 的研究就相当于对蛋白质骨架的研究，因此选择 C^α 的均方根偏差作为研究对象具有说服力。

均方根偏差是分子在模拟过程中运动幅度大小的一个衡量标准，是衡量体系是否稳定的重要依据，它显示了分子在模拟过程中的动态结构与天然态之间的相似性。RMSD 的值越大，表明在进行分子动力学过程中分子中的原子位置偏离初始构象的幅度越大，变化越剧烈；反之则表明分子位置的变化比较缓慢，分子的构象与初始结构相比变化趋势不是很明显。同时我们也可以通过 RMSD 的变化趋势来判定体系是否达到了平衡，当 RMSD 在某一范围内波动时就可以认为该体系达到了平衡^[68]。本实验选取 1cgd 能量优化后的得到的分子结构作为初始构象，选择 1cgd 的 C^α 的 RMSD 作为计算对象。

(2) 回转半径(Radius of Gyration, Rg)

回转半径是指蛋白质分子中所有原子与该蛋白质分子质量中心距离的几何平均值，它的计算公式如下：

$$R_g = \sqrt{\frac{1}{n} \sum_{i=1}^n (r_i - r_{cm})^2} \quad \text{式(3-5)}$$

式中: n - 蛋白质分子中原子的数目;

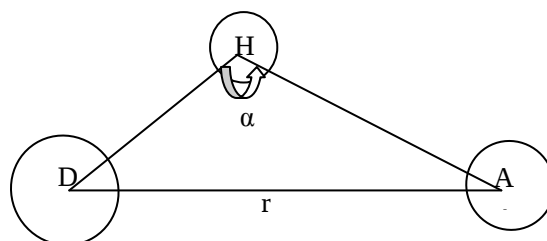
$r_i - r_{cm}$ - 原子 i 相对于该蛋白质分子质量中心的距离

R_g 一般用来衡量蛋白质分子结构的紧密程度, 对于具有相同原子数目的蛋白质分子, 回旋半径 R_g 反映了原子结合的紧密程度: R_g 越小, 则原子结合得越紧密。

(3) 氢键(Hydrogen Bond, hb)

氢键是指在氢原子与电负性很大的原子D以共价键结合时与另外一个具有孤对电子对的原子A之间形成的一个键, 是一种由氢原子参加成键的特殊化学键, 分为分子间氢键和分子内氢键两大类。它们之间的结合是以H离子为桥梁而形成的, 故称为氢键, 它的通式表达式为: $-D-H \cdots A-$, 氢键形成时, 提供氢原子的原子D叫做氢供体原子, 与氢原子配位的原子A叫氢受体原子, 其中D-H键一般为共价键, $H \cdots A$ 键为较弱的有方向性的范德华引力。

本实验中我们评判氢键存在的几何指标为: 氢供体与氢受体之间的距离不大于0.3nm, 而且氢供体-氢原子-氢受体之间的夹角不小于 120° 时我们认为有氢键形成。如图3.6为氢键示意图。蛋白质分子内氢键的数目越多, 分子内静电相互作用就越大, 蛋白质的稳定性就越高。



D - 氢供体原子 H- 氢原子 A- 氢受体原子

α - 供体-氢原子-受体之间的夹角, $\geq 120^\circ$

r - 氢供体与氢受体之间的距离, $\leq 0.3\text{nm}$

图 3.6 氢键示意图

Fig 3.6 The schematic diagram of hydrogenbonds

从这三个参数的简介我们知道这些参数能够很好地反映蛋白质分子某一个方面的构象状态的变化, 例如回转半径反映了蛋白质分子结构的松散或者紧密, 氢键反映了蛋白质分子内静电相互作用的大小, 均方根偏差虽然能够反映蛋白质分子结构是否稳定, 但是它容易受肽链活性末端的影响而产生较大的变化, 因此单独作为评价稳定性的指标

也不具有足够的说服力。所以本实验选择RMSD, Rg和hb三个参数综合说明温度的变化对1cgd I型胶原蛋白结构的影响。

3.4.2 模拟轨迹的分析过程

前两个参数, RMSD 和 Rg 的分析是用 Compaq Visual.Fortran.Version.6.6 编写的 order_p.exe 程序, 出于保密协议该程序不能在文中公开。首先新建一个文件夹, 把前一步不同温度条件下模拟得到的 1cgd.arc 文件、1cgd.seq 文件、1cgd.key 文件、1cgd.xyz_2 文件、amber99.prm 文件和编写的 order_p.exe 程序放入这个新建的文件夹中, 这些就是进行 RMSD 和 Rg 分析时所需要的文件。因为要对比的结构是优化后的结构, 因此选择了 1cgd.xyz_2 文件而不是 1cgd.xyz 文件, 在运行之前需要把 1cgd.xyz_2 文件重命名为 1cgd_native.xyz, 这是由 order_p.exe 程序所决定的。order_p.exe 程序运行完后得到的是 1cgd_instrand.txt 文件, 该文件是对 1cgd.arc 文件中的 200 个结构的轨迹文件与优化后的结构进行比较的量化分析, 我们选取 300K 时得到的该文件的前 5 行进行说明, 如表 3.2 所示。

表 3.2 1cgd_instrand.txt 文件的部分说明
Table 3.2 Section describes of 1cgd_instrand.txt

i	RMSD (Å)			Gyration (Å)	contact number	energy
	alpha	backbone	all			
1	1.231704	1.258051	1.408535	24.493293	204.000000	1171.584677
2	1.493176	1.508792	1.739025	25.052088	193.000000	1157.940591
3	2.146512	2.158628	2.417963	23.961931	191.000000	1190.142402
4	1.486564	1.512618	1.743455	24.819400	199.000000	1084.068386
5	1.045146	1.035171	1.213430	24.620747	203.000000	1130.617050

说明: i-序列号; rmsdalpha-C α 的 RMSD; rmsdbackbone-重原子的 RMSD;
rmsdall-全原子的 RMSD; gyration-回转半径; contact number-接触数; energy-能量

本实验要分析的为 C α 的 RMSD 和回转半径, 也就是 1cgd_instrand.txt 文件的第 2 列和第 5 列。要分析 RMSD 和 Rg 的话就需要绘制包含 200 个点的图形, 比较困难, 因此需要用到 MATLAB 软件进行图形的绘制分析。在 MATLAB 中编写了 g_nc, 如下:

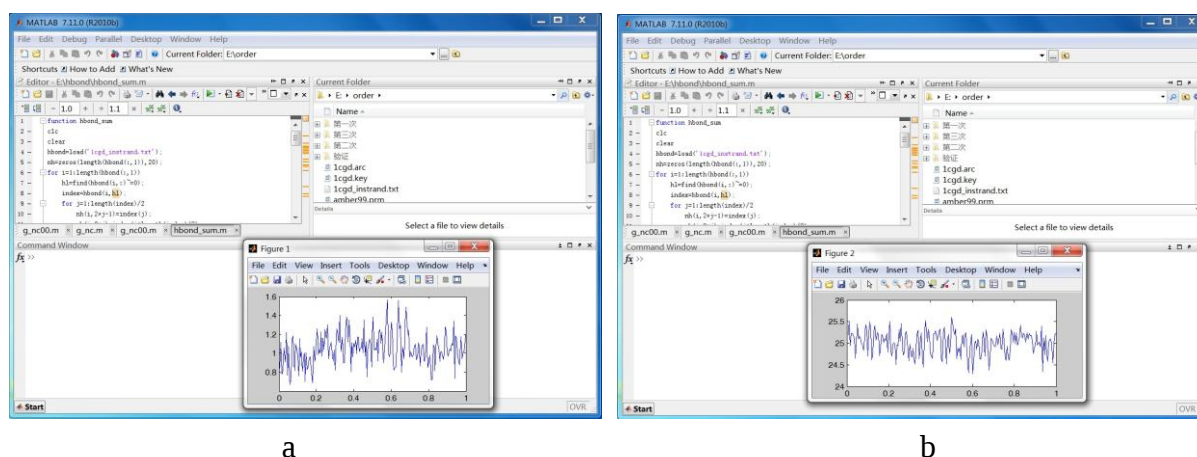
```
function g_nc
clc
clear
x=load('1cgd_instrand.txt');
```

```

t=(1:200)*5/1000;
figure(1)
plot(t,x(:,2))
figure(2)
plot(t,x(:,5))
end

```

用 x 代表从 1cgd_instrand.txt 里面提取数字； t 代表单位的换算，因为在模拟过程中是每隔 5ps 将体系中各个原子的坐标信息写入文件中，我们这里的时间是用 ns 作单位，因此需要进行单位的换算；figure（）表示选取哪一种选择，本实验是要分析 RMSD 和 Rg，因此要选取 1cgd_instrand.txt 中的第 2 列和第 5 列作为 Y 轴（单位：Å，也就是 Angstrom），以时间 t 作为 X 轴（单位：ns）。则得到的 figure 1 为 RMSD 随时间的变化情况，figure 2 为 Rg 随时间的变化情况，如图 3.7 所示。

a - 1cgd 在 MATLAB 中的 C^α RMSD 截图

b - 1cgd 在 MATLAB 中的 Rg 截图

图 3.7 1cgd 在 MATLAB 中的曲线截图

Fig 3.7 A curve screenshot of 1cgd in MATLAB

第三个参数 hb 的分析我们是用 Compaq Visual.Fortran.Version.6.6 编写了 hbond.exe 程序，出于保密协议该程序不能在文中公开。重新建一个文件夹，把前一步不同温度条件下模拟得到的 1cgd.arc 文件、1cgd.seq 文件、1cgd.key 文件、1cgd.xyz_2 文件、amber99.prm 文件和编写的 hbond.exe 程序放入这个新建的文件夹中，这些就是进行氢键分析时所需要的文件。因为要得到的氢键数目是优化后的蛋白质的氢键数目，因此我们选择了 1cgd.xyz_2 文件而不是 1cgd.xyz 文件，在运行之前需要把 1cgd.xyz_2 文件重命名为 1cgd_native.xyz，这是由 hbond.exe 程序所决定的。运行完后得到的同样是名称为 1cgd_instrand.txt 文件，为了避免把分析 RMSD 和 Rg 时得到的 1cgd_instrand.txt 文件混

淆和覆盖我们才会在分析前重新建文件夹。同样，由于得到的数据过多，为了便于统计，在 MATLAB 中编写了 hbond_sum 用于统计氢键的种类和数目，该程序见附录 I。

3.5 本章小结

本研究使用 TINKER 5.1 软件在正则系综下，采用 GB/SA 隐含水模型和 AMBER 99 力场对 1cgd 这种 I 型胶原蛋白在 260K，300K，340K 和 380K 四个温度下进行分子动力学模拟，得到的轨迹文件用 Compaq Visual Fortran Version. 6.6 编写的 order_p.exe 程序进行 RMSD 和 Rg 的分析，hbond.exe 程序进行氢键的分析，为了使实验结果便于分析又采用 MATLAB R2011b 软件编写了 g_nc 程序绘制 RMSD 和 Rg 的曲线图以及 hbond_sum 程序对氢键的种类和数目进行统计。实验过程如图 3.8 所示。

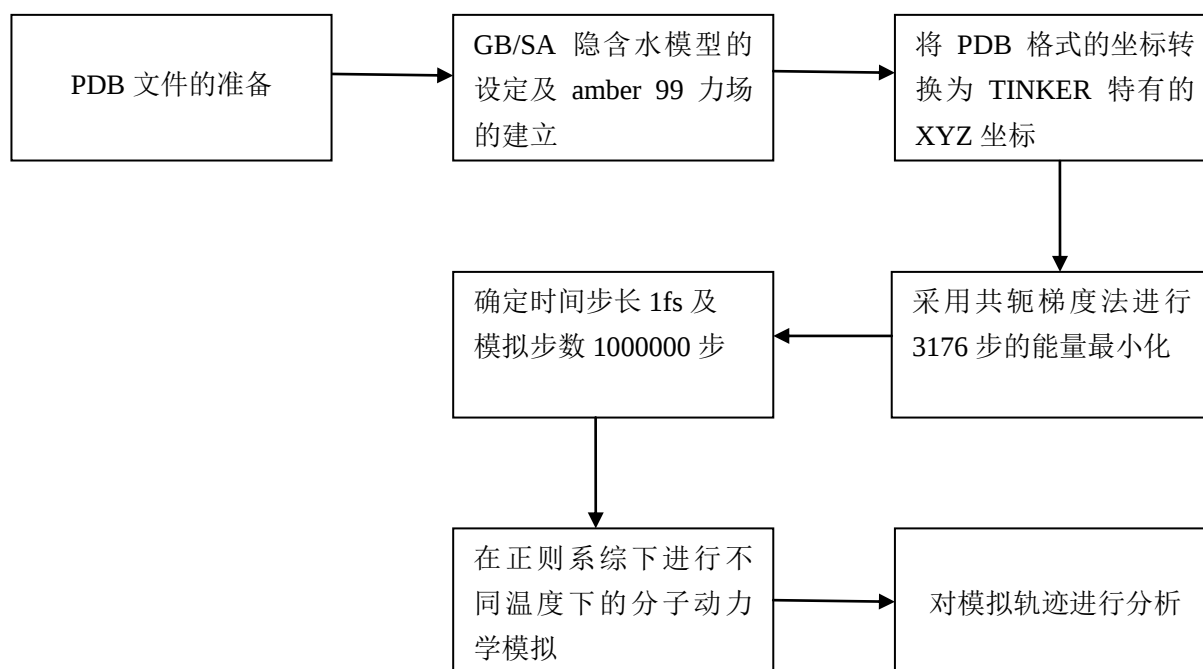


图 3.8 实验过程流程图

Fig 3.8 The experimental process flow chart

4 实验结果的分析与讨论

通过对第三章实验部分得到的 C^α 的 RMSD、所有原子的 Rg 和氢键的分析, 以及结合 Force Field Explorer 5.0 软件所观察到的 1cgd 在不同温度条件模拟下的空间构象随时间变化所得到的图形, 分析 1cgd 这种 I 型胶原蛋白在不同温度条件下的稳定性, 从而从微观角度去了解天然皮革的热稳定性。

为了确保实验结果的准确性, 减少计算机运行时所造成的不可避免的误差, 对每个温度下的分子动力学模拟都重复进行 5 次, 最后实验结果的分析取这五次模拟结果的平均值。

4.1 RMSD 分析

前面已经介绍了 RMSD 反应的是蛋白质分子在进行分子动力学模拟的过程中产生的结构相对于优化后的初始结构的偏离程度。图 4.1 中的(a)、(b)、(c)、(d)分别为 I 型胶原蛋白 1cgd 在 260K、300K、340K 和 380K 的 C^α RMSD 值的曲线变化图。

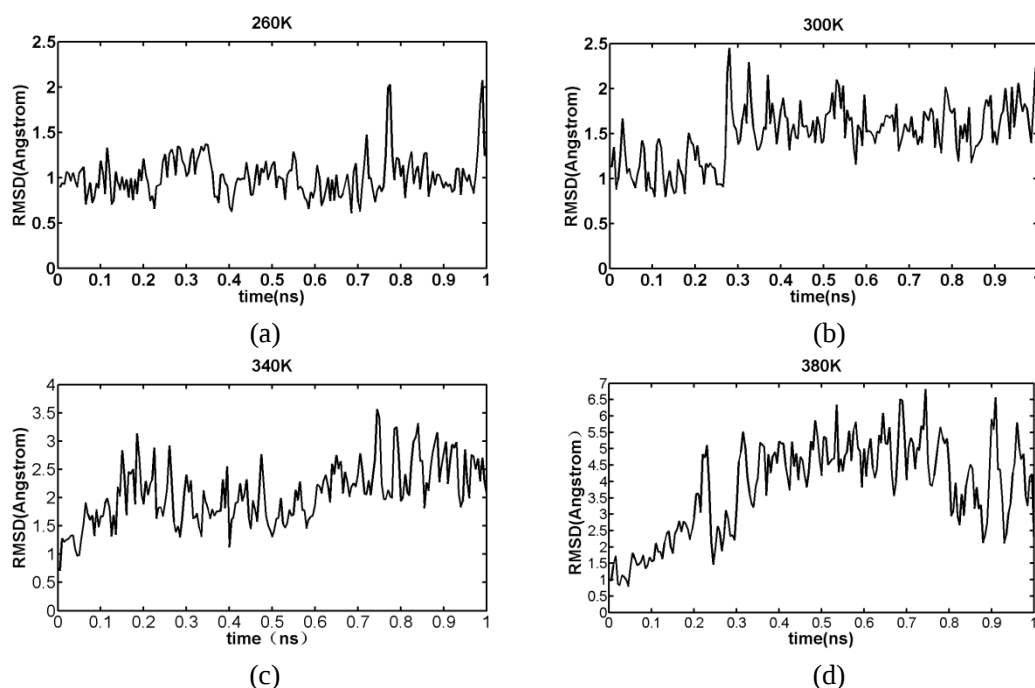


图 4.1 不同温度的 1cgd 的 C^α RMSD 图

Fig 4.1 The alpha RMSD values of 1cgd at different temperature

说明: (a) 260K 时 1cgd 的 C^α RMSD 图

(b) 300K 时 1cgd 的 C^α RMSD 图

(c) 340K 时 1cgd 的 C^α RMSD 图

(d) 380K 时 1cgd 的 C^α RMSD 图

从图 4.1 (a) 中可以看出在 260K 的温度下, 1cgd 的 RMSD 值变化不大, 基本维持在 1Å 左右, 没有太大的波动, 这就说明此时 1cgd 的结构接近于天然态的结构, 显示了比较高的稳定性。

图 4.1 (b) 中, 1cgd 在温度为 300K 时, 前 0.25ns 它的 RMSD 值基本在 1.1Å 附近波动, 而 0.25ns 后 RMSD 值突然上升到接近于 2.5Å, 0.4ns 后 RMSD 值基本维持在 1.6 Å 左右, 此时体系达到平衡。从 RMSD 值得变化情况可以看出, 在 0.25ns 之前 1cgd 的结构接近于天然构象, 0.25ns 以后才看出 1cgd 的结构发生了变化, 这可能是由于在进行分子动力学模拟时系统的温度升高到指定温度有一个过程需要一定的时间, 因此模拟初期结构没什么变化。当系统达到稳定时, RMSD 值大约为 1.6Å, 说明 1cgd 在 300K 时虽然结构发生了一些改变但仍然比较稳定。

图 4.1 (c) 所显示的 1cgd 的 RMSD 值在 340K 的情况下变化的就没有 1cgd 在温度为 260K 和 300K 时那么规律。从 0 到 1.5ns, RMSD 值呈上升趋势, 从 0.7Å 增大到 2Å 左右, 在 1.5ns-0.5ns 这段时间 RMSD 值在 2Å 附近波动, 而从 0.5ns 开始, RMSD 值又一次的变大, 直到 0.75ns 才趋于平稳, 大约在 2.2Å。说明 1cgd 在 340K 时的结构与天然态的相比已经有较大的变化, 变得不太稳定。

从图 4.1 (d) 中看出, 当温度为 380K 时 1cgd 的 RMSD 值在模拟初期就上升, 从 1Å 上升到 5Å, 涨幅很大, 虽然在 0.75ns 以后有一些下降, 但是仍然在较大的范围内波动。RMSD 值很大, 说明 1cgd 在 380K 时的结构与天然态的结构相比, 变化程度很大, 1cgd 很不稳定, 虽然在 0.75ns 以后 RMSD 值有所下降, 但是波动性仍然很大, 说明 380K 时 1cgd 的分子构象没有达到相对平衡的状态, 需要进行更长时间的分子动力学模拟来进一步的观察。

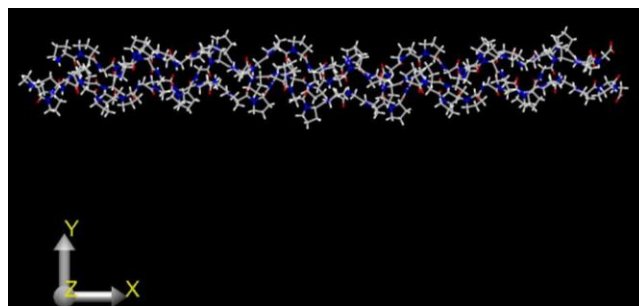
通过对图 4.1 (a) - (d) 四幅图的观察和分析我们可以总结出 1cgd 在温度升高的情况下, RMSD 值有增大的趋势, 也就是说, 当温度升高时 1cgd 的结构与天然态的结构相比, 变化的程度会增大。说明温度对 1cgd 结构的稳定性有影响, 温度升高, 稳定性下降, 特别是当温度超过 340K 以后, 1cgd 变得很不稳定。

4.2 轨迹取样分析

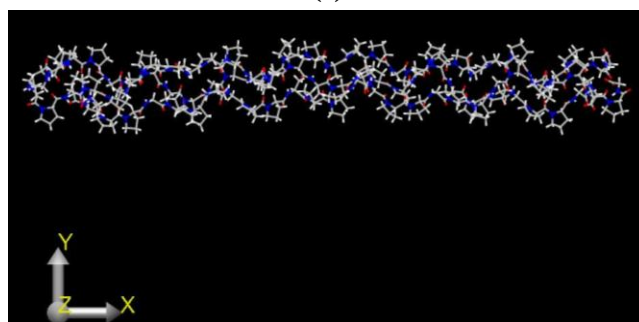
前面分析了当模拟的温度升高时, RMSD 值变大, 1cgd 的结构偏离天然态结构的程度也就越大, 为了更好的理解这一变化, 我们采用 Force Field Explorer 5.0 软件对 1cgd 在模拟过程中 RMSD 值变化显著的运动轨迹进行取样分析, 这样就能把比较抽象的结果以图形的方式更加直观的呈现出来, 方便理解 1cgd 的三维空间结构随模拟温度的改变而不断变化的情况。图 4.2、4.3、4.4 和 4.5 分别为 1cgd 在温度为 260K、300K、340K 和 380K 的空间构象的轨迹取样图。

4.2.1 260K 轨迹取样分析

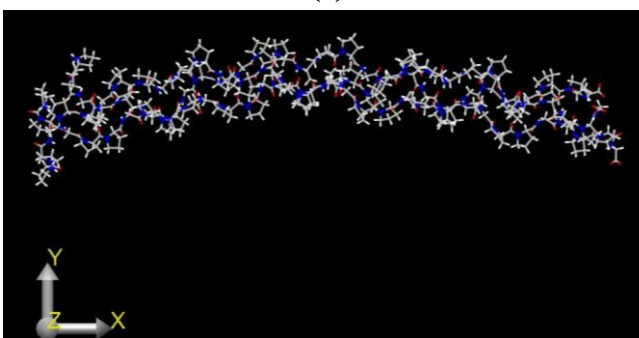
260K 时 1cgd 的 RMSD 值变化不大, 我们截取了 0.5ns 和 0.9ns 时刻的轨迹图, 也就是空间构象图进行分析, 把它们与 1cgd 天然态的空间构象进行对比分析, 如图 4.2 所示。



(a)



(b)



(c)

图 4.2 1cgd 在 260K 时的轨迹取样图

Fig 4.2 The trajectories of 1cgd at 260K

说明: (a) 1cgd 的天然态构象

(b) 1cgd 在 260K 温度下 0.5ns 时刻的轨迹取样图

(c) 1cgd 在 260K 温度下 0.9ns 时刻的轨迹取样图

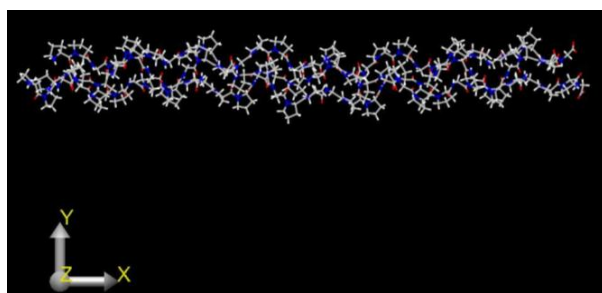
由于 RMSD 分析的是 1cgd 在不同时刻的结构与天然态结构的偏差程度, 所以我们

截取了 1cgd 的天然构象作为参考, 也就是图 4.2(a), 从图中可以看出, 1cgd 在天然态的时候是由 3 条 α 螺旋缠绕在一起的超螺旋结构。该图像中还包含了 X/Y/Z 全球轴, 图中所显示的全球轴方向即是我们观察 1cgd 结构的方向, 为了更好地说明结构的变化情况, 我们所有的轨迹取样图都采用该方向, 这在其余的轨迹取样图中也有显示。

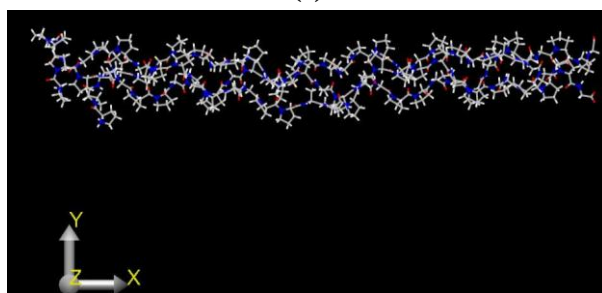
将 0.5ns 和 0.9ns 时刻的空间构象与 1cgd 的天然构象进行比较, 发现在 260K 温度下, 1cgd 的空间构象变化不大, 仅仅是分子两端残基的位置发生了一些改变, 与前面分析的 RMSD 的变化情况相符。

4.2.2 300K 轨迹取样分析

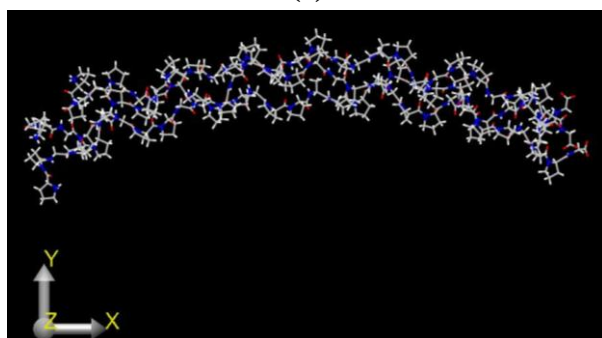
根据前面对 1cgd 的 RMSD 在 300K 温度下的分析, 我们知道在 0.25ns 以后 RMSD 的值有了较大的变化, 所以我们截取了 0.25ns, 0.55ns 和 0.95ns 时刻的轨迹图进行对比分析, 如图 4.3 所示。



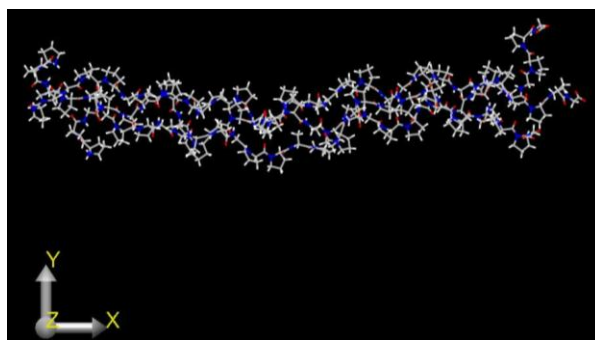
(a)



(b)



(c)



(d)

图 4.3 1cgd 在 300K 时的轨迹取样图

Fig 4.3 The trajectories of 1cgd at 300K

说明: (a) 1cgd 的天然态构象

(b) 1cgd 在 300K 温度下 0.25ns 时刻的轨迹取样图

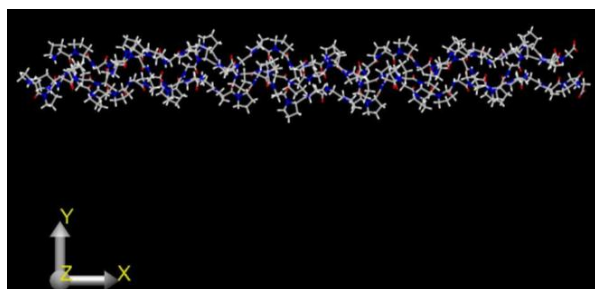
(c) 1cgd 在 300K 温度下 0.55ns 时刻的轨迹取样图

(d) 1cgd 在 300K 温度下 0.95ns 时刻的轨迹取样图

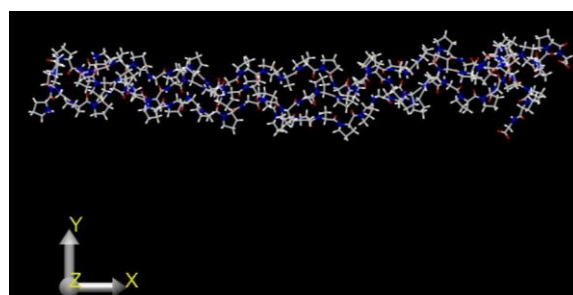
通过将图 4.3 中的 (b) 与 (a) 比较可以看出, 1cgd 的在 0.25ns 时刻的空间构象与天然态的相比, 变化不大, 但图 (c) 和 (d) 相对于图 (a) 而言结构的变化就比较明显了, 尤其是分子两端, 变得比较弯曲, 但是 (c) 和 (d) 之间的变化就很小, 说明 1cgd 的空间构象已经达到了相对平衡。这些轨迹取样的结果也证实了 1cgd 在 300K 温度下 RMSD 值的变化趋势, 在 0.25ns 以前变化不大, 0.25-1ns 之间 RMSD 值的变化增大, 但处于一个相对稳定的范围内波动。

4.2.3 340K 轨迹取样分析

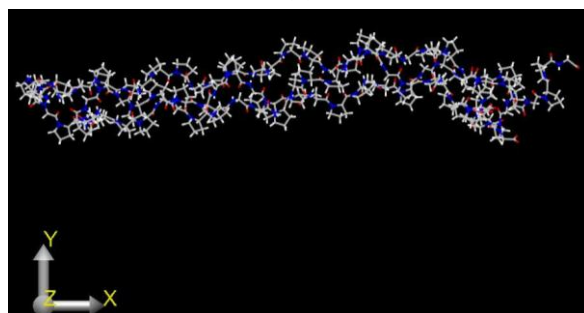
从前面的轨迹取样分析我们已经知道, 当 RMSD 值比较小的时候空间构象与天然态相比没有太大的变化, 所以 340K 温度下我们只截取了 1cgd 结构变化后趋于稳定的空间构象, 所以我们截取了 0.8ns, 0.9ns 和 1.0ns 时刻的轨迹图, 如图 4.4 所示。



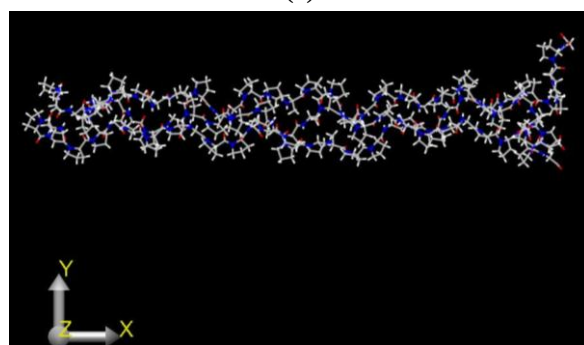
(a)



(b)



(c)



(d)

图 4.4 1cgd 在 340K 时的轨迹取样图

Fig 4.4 The trajectories of 1cgd at 340K

说明: (a) 1cgd 的天然态构象

(b) 1cgd 在 340K 温度下 0.8ns 时刻的轨迹取样图

(c) 1cgd 在 340K 温度下 0.9ns 时刻的轨迹取样图

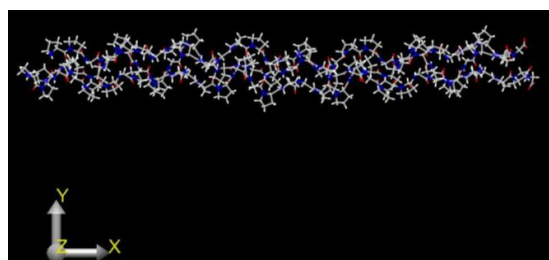
(d) 1cgd 在 340K 温度下 1.0ns 时刻的轨迹取样图

通过这四幅图片的观察与对比,我们发现 1cgd 在 340K 温度下的空间构象相对于天然态变化比较大,天然态时的三条 α 螺旋已经变得不那么规则了,但这三个时刻 1cgd 的结构没有什么太大的变化,说明 1cgd 的空间构象已经达到了相对平衡。

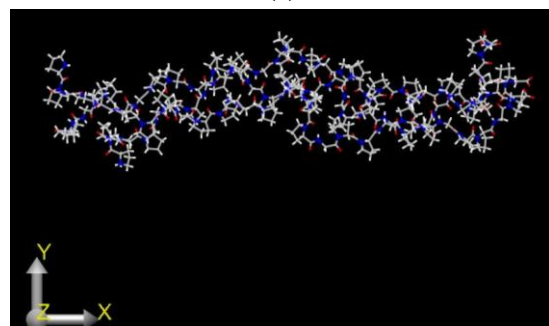
4.2.4 380K 轨迹取样分析

380K 时 1cgd 的 RMSD 曲线图显示出该温度下的空间构象变化很大,我们截取

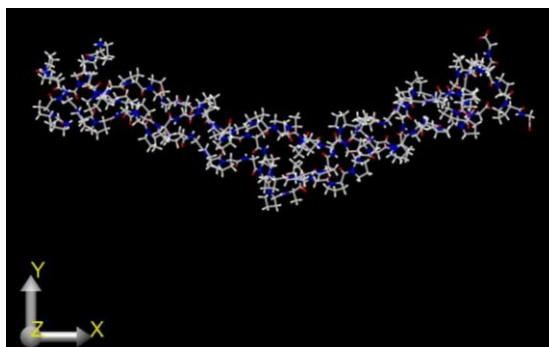
0.4ns, 0.6ns 和 0.9ns 时刻的轨迹进行取样分析, 如图 4.5 所示。



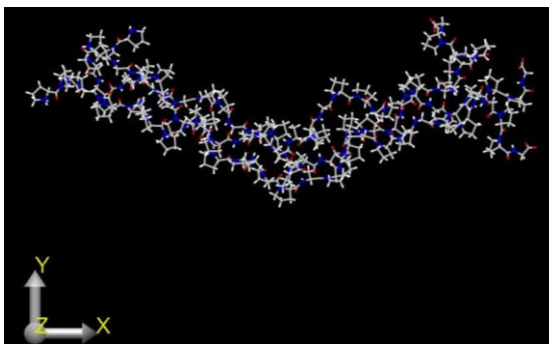
(a)



(b)



(c)



(d)

图 4.5 1cgd 在 380K 时的轨迹取样图

Fig 4.5 The trajectories of 1cgd at 380K

说明: (a) 1cgd 的天然态构象

(b) 1cgd 在 380K 温度下 0.4ns 时刻的轨迹取样图

(c) 1cgd 在 380K 温度下 0.6ns 时刻的轨迹取样图

(d) 1cgd 在 380K 温度下 0.9ns 时刻的轨迹取样图

图 4.5 展示了 1cgd 在 380K 温度下构象的剧烈变化情况，由天然态的类似杆状或者棒状变成了向中间折叠的弯曲状态，变化非常显著，而且在整个模拟的过程中构象一直在发生改变而没有达到过相对平衡的状态，这跟前面分析 RMSD 的结果相同。

4.3 Rg 分析

Rg 是指蛋白质分子中所有原子与该蛋白质分子质量中心距离的几何平均值，它反映了蛋白质分子的松散或者紧密的程度，图 4.6 显示出 1cgd 这种 I 型胶原蛋白在 260K、300K、340K 和 380K 这四种温度下回转半径的变化情况。

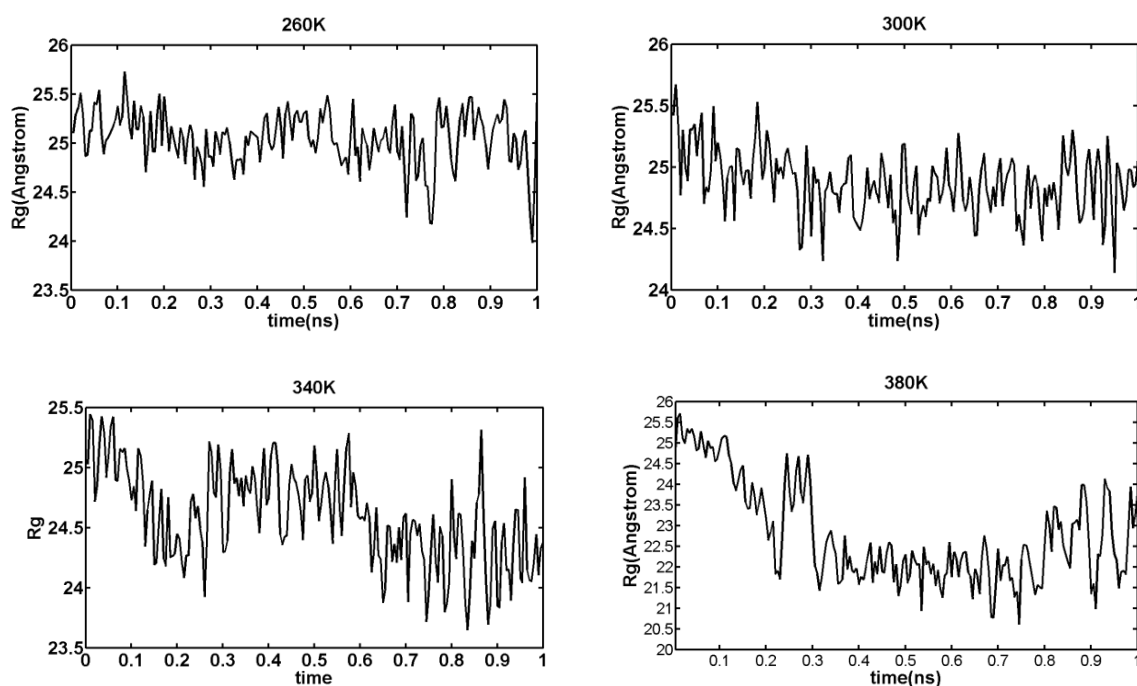


图 4.6 1cgd 在不同温度下的 Rg 变化图

Fig 4.6 The Rg values of 1cgd at different temperature

从图中可以看出 1cgd 在 260K 和 300K 这两种温度下，Rg 值比较稳定，分别在 25Å 和 24.8Å 附近波动，当温度上升为 340K 时，Rg 值出现了波动，从 25.1Å 下降到 24.4Å，再上升为 24.8Å，最后在 24.3Å 附近波动。在 380K 时，Rg 值的波动就更剧烈了，从 25.2Å 直降到 21.8Å，然后迅速上升为 24.8Å，再几经波动大概在 22.2Å 附近波动，在 0.8ns 以后又出现了波动，这可能也跟前面分析 RMSD 的原因一样，由于模拟的时间太短，1cgd

的构象在 380K 温度下还没有完全达到相对平衡。

观察 1cgd 的 R_g 值从 260K 到 380K 的整体变化情况,发现当温度升高时 R_g 值下降,从 25Å 下降到 22.2Å。这就说明 1cgd 蛋白质分子的结构随着温度的上升而变得比较紧密。

4.4 氢键分析

分析氢键时是在 MATLAB 中编写了 hbond_sum.m 用于统计氢键的种类和数目,得到的结果中有些种类的氢键数目很少,比如只有 1 个,为了分析的方便,我们把氢键数目占氢键总数百分比 < 3% 的氢键类型就不列举在表中进行分析。

表 4.1、4.2、4.3 和 4.4 列举出了 1cgd 在温度分别为 260K、300K、340K 和 380K 时, 0-1ns 这段模拟时间内模拟轨迹中氢键的出现情况。

表 4.1 260K 时 1cgd 形成氢键情况 (>3%)
Table 4.1 The hydrogenbonds of 1cgd at 260K (>3%)

氢键类型	氢键数目 (个)	所占比率
4-66	189	19.8%
37-5	184	19.3%
4-65	182	19.1%
34-3	156	16.3%
34-2	99	10.4%
35-3	66	6.9%
5-66	31	3.2%
共: 16 种	共: 955 个	

首先对表中的氢键类型进行加以解释说明: 前面介绍 PDB 文件的准备时已经指出 1cgd 包括 A、B、C 三条链, 每条链上 30 个残基, 而且每条链上残基的序列都相同, 我们对 A、B、C 三条链整体进行编号, 给 A 链上的第一个残基编号为“1”, 其余的残基按顺序依次编号, 那么 1-30 号就表示 A 链上的残基, 31-60 表示 B 链上的残基, 61-90 就表示 C 链上的残基, 而且每个编号所对应的残基类型可以根据每条链上的残基序列得知, 例如氢键类型“4-66”就表示 A 链上的 4 号残基和 C 链上的 6 号残基之间形成的氢键, 对照每条链上的残基序列可知, 也就是 A 链上的 4 号脯氨酸和 C 链上的 6 号甘氨酸之间形成的氢键, 其余的氢键类型以此类推。

从表 4.1 可以看出, 260K 时 1cgd I 型胶原蛋白在 0-1ns 这段模拟时间内总共出现了 16 种不同类型的氢键, 氢键总数为 955 个, 观察氢键类型发现只存在不同链间形成的氢键, 例如 A 链与 C 链之间形成的氢键, 或者 A 链与 B 链之间形成的氢键, 所占比率超

过 10% 的氢键类型为 4-66, 37-5, 4-65, 34-3 和 34-2。

表 4.2 300K 时 1cgd 形成氢键情况 (>3%)
Table 4.2 The hydrogenbonds of 1cgd at 300K (>3%)

氢键类型	氢键数目 (个)	所占比率
4-66	182	20.0%
37-5	181	19.8%
4-65	172	18.9%
34-3	111	12.1%
34-2	92	10.1%
2-63	39	4.3%
2-62	35	3.8%
5-66	30	3.3%
共: 17 种	共: 912 个	

300K 时 1cgd I 型胶原蛋白在 0-1ns 这段模拟时间内总共出现了 17 种不同类型的氢键, 氢键总数为 912 个, 观察氢键类型发现只存在 A 链与 C 链之间形成的氢键, 或者 A 链与 B 链之间形成的氢键, 所占比率超过 10% 的氢键类型为 4-66, 37-5, 4-65, 34-3 和 34-2。

表 4.3 340K 时 1cgd 形成氢键情况 (>3%)
Table 4.3 The hydrogenbonds of 1cgd at 340K (>3%)

氢键类型	氢键数目 (个)	所占比率
37-5	174	20.5%
4-65	166	19.6%
4-66	156	18.4%
34-3	103	12.1%
34-2	94	11.8%
35-3	43	5.1%
5-66	39	4.6%
共: 21 种	共: 849 个	

340K 时 1cgd I 型胶原蛋白在 0-1ns 这段模拟时间内总共出现了 21 种不同类型的氢键, 氢键总数为 849 个, 观察氢键类型发现也只存在不同链间的氢键, 所占比率超过 10% 的氢键类型为 37-5, 4-65, 4-66, 34-3 和 34-2。

380K 时 1cgd I 型胶原蛋白在 0-1ns 这段模拟时间内总共出现了 18 种不同类型的氢键, 氢键总数为 735 个, 氢键类型仍然是只存在不同链间的氢键, 所占比率超过 10% 的氢键类型为 37-5, 4-65, 4-66, 2-63 和 2-62。

表 4.4 380K 时 1cgd 形成氢键情况 (>3%)
Table 4.4 The hydrogenbonds of 1cgd at 380K (>3%)

氢键类型	氢键数目	所占比率
37-5	179	24.4%
4-65	162	22.0%
4-66	156	21.2%
2-63	90	12.2%
2-62	80	10.9%
5-66	46	6.3%
共: 18 种	共: 735 个	

结合表 4.1-4.4, 我们整体分析一下 260K、300K、340K 和 380K 四个温度下 1cgd I 型胶原蛋白在 0-1ns 这段模拟时间内氢键的情况:

首先从氢键的数目上分析, 在 0-1ns 的整个模拟过程中, 1cgd 在温度为 260K 时总共出现了 955 个氢键, 300K 时出现过 912 个氢键, 340K 时 849 个氢键, 而到了 380K 时出现的氢键个数为 735 个, 也就是说 1cgd 这种 I 型胶原蛋白所含的氢键数目随着温度的上升而减少, 而且温度越高氢键数目减少的越快。

从氢键类型上分析, 不管是以上四个表格中出现的氢键类型, 还是所占比率小于 3% 而没有列举出来的氢键类型, 我们发现在 0-1ns 的整个模拟过程中出现的氢键类型只有链与链之间形成的氢键, 比如 A 链与 C 链之间形成的氢键, 或者 A 链与 B 链之间形成的氢键, 而没有发现链内形成的氢键, 也就是说, 1cgd I 型胶原蛋白中只存在分子间氢键, 而不存在分子内氢键。

从氢键的持续性方面分析, 表 4.5 列举了 1cgd I 型胶原蛋白处于不同温度时在 0-1ns 的整个模拟时间内, 持续性超过 5% 的氢键类型在总的氢键类型中所占的比率。

表 4.5 不同温度下氢键的持续性
Table 4.5 The persistence of hydrogenbonds at different temperature

温度 (K)	总的氢键种类	所占比率>5%的氢键种类	所占比率 (%)
260	16	6	37.5
300	17	5	29.4
340	21	6	28.6
380	18	6	33.3

从表 4.5 中我们知道在 260K, 300K, 340K 和 380K 温度下 1cgd I 型胶原蛋白中氢键持续性>5%的氢键类型在总的氢键类型中所占的比率分别为 37.5%, 29.4%, 28.6% 和 33.3%, 这些数据表明: 1cgd I 型胶原蛋白在 0-1ns 的模拟过程中超过半数的氢键类型所占比率不超过 5%, 也就是大部分氢键在模拟过程中的持续性很低, 表明 1cgd I 型

胶原蛋白中大部分类型的氢键稳定性很差。

整体比较表 4.1-4.4 我们发现，稳定性比较好的氢键主要是以下三种类型的氢键：4-66、37-5 和 4-65，也就是 A 链上的 4 号脯氨酸和 C 链上的 6 号甘氨酸之间形成的氢键，B 链上的 7 号脯氨酸和 A 链上的 5 号羟脯氨酸之间形成的氢键，以及 A 链上的 4 号脯氨酸和 C 链上的 5 号羟脯氨酸之间形成的氢键。虽然有的氢键在某一温度下所占全部氢键的比率也比较高，例如 34-3，即 B 链上的 4 号脯氨酸和 A 链上的 3 号甘氨酸之间形成的氢键，在 260K 时所占的比率为 16.3%，在 300K 和 340K 时都为 12.1%，但是当温度为 380K 时，该类型的氢键所占比率不到 3%，所以整体上说这类氢键的稳定性一般。

4.5 本章小结

(1) 通过分子动力学模拟的方法建立了 1cgd I 型胶原蛋白在不同温度下的分子模型，并且其最终热稳定性的结果显示与天然皮革的宏观性能是相吻合的，说明该模型是成功的。选择 I 型胶原蛋白作为皮革热稳定性的研究对象是因为大量的文献资料显示 I 型胶原蛋白在不同成年动物皮中所占的比例是最高的，因此研究它是研究天然皮革微观结构的重要手段。本实验选择 TINKER 软件进行分子动力学模拟，因为它具有很多优势，例如软件免费，而且操作比较简单，对于非专业人士也可以操作使用，对电脑设备的要求不高等，但是该软件仍在开发中，某些方面还不完善，需要我们进一步的调试。

(2) RMSD 值的曲线变化图以及对应时刻的轨迹截取图表明，温度升高时 I 型胶原蛋白稳定性会下降，结构会发生变化，而且温度越高结构变化越明显，如 RMSD 值从 260K 的 1Å 变化到 380K 的 5Å 时，空间构象也由天然态的类似杆状或者棒状变成了向中间折叠的弯曲状态，变化非常明显。

(3) Rg 值从 260K 的 25Å 下降到 380K 的 22.2Å，这种变化表明 I 型胶原蛋白随着温度的上升结构变得比较紧密，这一结论或许可以解释皮革高温时发生收缩的现象，不过需要进一步的实验来验证。

(4) 分子内氢键是维持 I 型胶原蛋白稳定性的主要作用力，当氢键数目由 260K 时的 955 个减少到 380K 时的 735 个时，分子内相互作用的强度下降导致 I 型胶原蛋白的热稳定性下降。

(5) I 型胶原蛋白中大部分种类的氢键稳定性很差，比较稳定的氢键类型包括三种：A 链上的 4 号脯氨酸和 C 链上的 6 号甘氨酸之间形成的氢键；B 链上的 7 号脯氨酸和 A 链上的 5 号羟脯氨酸之间形成的氢键；A 链上的 4 号脯氨酸和 C 链上的 5 号羟脯氨酸之间形成的氢键。这一结论对于提高皮革的热稳定性有一定的指导作用。

致 谢

在此论文完成之际，我要向帮助我完成论文的所有人表示衷心的感谢和诚挚的敬意！

首先我要向我的导师钟安华教授表示最由衷的感谢。钟老师在科研上敏锐的思维，严谨的科研态度，渊博的学识深深的影响着我，生活中果断干练的作风，诚挚谦虚的品格和宽厚善良的处事方式深深的感染着我。在整个三年的学习阶段特别是论文的撰写过程中倾注了大量的心血，从选题、资料搜集到论文写作阶段，钟老师都会抽出时间耐心的给予指导，论文初稿完成后又不厌其烦的一次次批改，提出许多中肯的指导意见，使我在研究和写作过程中始终把握住方向。生活中钟老师也如慈母般关心和爱护我。在此我想发自肺腑的说一句：“钟妈，谢谢您，我会更加努力的！”

本论文中技术性的问题，如各种软件的使用以及编程都离不开服装学院江学为博士的耐心教授和悉心指导，江博士渊博的知识和精湛的编程技艺让我深刻体会了知识就是力量这句话，在此表示我衷心的感谢！

感谢我的同学及室友许维，张四华和张玲，感谢你们在生活和学习上对我的帮助和支持，尤其感谢许维同学在论文完成过程中提出的宝贵意见和建议！

另外感谢我的家人和朋友对我的宽容、理解、支持和鼓励，让我在遇到困难时仍然奋勇向前！

最后感谢服装学院的所有老师和领导对我在校七年的培养和帮助，感谢参加盲审和答辩的各位老师！

参考文献

- [1] 丁绍兰. 革制品材料学[M]. 北京: 中国轻工业出版社, 2001, 6-16
- [2] 施辉阳. 酸法提取生猪皮胶原蛋白的研究[D]. 北京: 北京工业大学, 2004
- [3] 蒋挺大, 张春萍. 胶原蛋白[M]. 北京: 化学工业出版社, 2001, 1-19
- [4] Pace J.M., Corrado M., et al. Identification, characterization and expression analysis of a new fibrillar collagen gene [J]. *Matrix Biology*, 2003, 22(1): 3-14
- [5] 王荣春, 徐德昌, 秦兰霞. 胶原蛋白与 V 型胶原蛋白的结构概述[J]. *生物信息*, 2004, 2(4): 30-31
- [6] 阎隆飞, 孙之荣. 蛋白质分子结构[M]. 北京: 清华大学出版社出版, 1999, 27-95
- [7] Ramachandran G N, Kartha G. Structure of collagen [J]. *Nature*, 1955, 176: 593-595
- [8] 王堃, 郑学晶, 孟卓君等. 影响胶原结构与性能的因素研究进展[J]. *化工进展*, 2010, 29(1): 88
- [9] Smith J W. Molecular pattern in native collagen [J]. *Nature*, 1968, 219(5150): 157-164
- [10] Meullenet JF. Textural properties of chicken frankfurters with added collagen fibers [J]. *Journal of Food Science*, 1994, 59(4): 729-733
- [11] 陈秀金, 曹健, 汤克勇. 胶原蛋白和明胶在食品中的应用[J]. *郑州工程学院学报*, 2002, 23(1): 66-69
- [12] Mattsson I, Lorentzen JC. Immunization with Alum-Collagen II complex suppresses the development of collagen Induced arthritis in rats by deviating the immune response [J]. *Scandinavian Journal of Immunology*, 1997, 46(6): 619-624
- [13] 李彦春, 祝德义. 胶原多肽钙的制备及小鼠应用试验[J]. *中国皮革*, 2005, 34(15): 36-40
- [14] Chvapil M, Eckmayer Z. Role of proteins in cosmetics [J]. *Inter. J. Cosmet. Sci*, 1985, (7): 41-49
- [15] Peng Y, Glattauer V, et al. Evaluation for collagen products for cosmetic application [J]. *International Journal of Cosmetic Science*, 2004, 26(6): 313-314
- [16] Young. 应用纳米技术的胶原蛋白肌肤还原法[J]. *美容院*, 2002, (1): 26-27
- [17] Giudici G, Viola M. Molecular stability of chemically modified collagen triple helices [J]. *Federation of European Biochemical Societies*, 2003, 547(3): 170-176
- [18] 王永胜, 侯春林, 陈爱民等. 胶原海绵止血功能的实验研究[J]. *中国修复重建外科杂志*, 2001, 15(3): 140-143
- [19] Judith Hohlfeld, et al. Tissue engineered fetal skin constructs for paediatric burns [J]. *The Lancet*, 2005, 366(9): 840-842

- [20] 付丽红, 张铭让, 齐永钦等. 胶原蛋白和植物纤维结合机理的研究[J]. 中国造纸学报, 2002, 17(1): 68-71
- [21] 贾鹏翔, 汤克勇. 胶原蛋白改性丙烯酸类复鞣剂的制备[J]. 精细化工, 2006, 23(8): 801-805
- [22] 王鸿儒. 皮革生产的理论与技术[M]. 北京: 中国轻工业出版社, 1999, 253-285
- [23] 黄育珍. 皮和革在不同湿含量下的热稳定性[J]. 皮革科学与工程, 1994, 4(3): 32-33
- [24] 贾鹏翔. 皮革中胶原的聚集态结构及热稳定性研究[D]. 河南: 郑州大学, 2006
- [25] 杨文华. 表示扫描量热法研究皮革和羊皮的湿热稳定性[J]. 西部皮革, 2011, 2(4): 53
- [26] S.Menashi, A.Finch, et al. Enthalpy changes associated with the denaturation of collagens of different amino acid content [J]. Biochem. Biophys. Acta, 1976, 144: 623-625
- [27] 陈新江, 马建中, 卢燕等. 加热介质对皮革收缩温度的影响[J]. 中国皮革, 2003, 32(17): 18-24
- [28] M. Komanowsky. Thermal stability of hide and leather at different moisture contents [J]. JALCA, 1991, 86(8): 269-280
- [29] Bowes J H, Carter C W. The interaction of aldehydes with collagen[J]. Biochim Biophys. Acta, 1968, 168(2): 341-352
- [30] 王伟, 马建中, 杨宗遂. 皮革鞣剂及鞣制机理综述[J]. 中国皮革, 1997, 26(8): 31
- [31] 王亚娟, 郭军, 单志华. 溶剂对不同革热稳定性的影响[J]. 中国皮革, 2009, 2(3): 22
- [32] K. Takenouchi, K. Konda, et al. Effect of masked chromium Complexes on thermal stability of collagen [J]. Soc.Leaner Technol.Chem, 1995, 79 (2): 188
- [33] 康俊霞, 康永锋, 包斌等. Na^+ 、 Ca^{2+} 和pH值对鲸鲨皮胶原蛋白热变性温度的影响[J]. 食品科学, 2011, 32(13): 69
- [34] Regina K P, Ruminan K. Thermal stability of calf skin collagen type I in salt solutions [J]. Biochimica et Biophysica Acta, 1996, 1297(2): 171-181
- [35] Chahine. Changes in hydrothermal stability of leather and parchment with deterioration: a DSC study [J]. Thermochemica Acta, 2000, 365: 101-110
- [36] 周鲁, 石碧. 皮革热收缩的热力学和动力学理论基础[J]. 中国皮革, 2003, 32(1): 15-17
- [37] A.D.Covington. Studies of the origin of hydrothermal stability: a new theory of tanning [J]. JALCA, 1998, 93(3): 107-120
- [38] 汤克勇, 王芳等. 鞣制对皮革热收缩行为的影响[J]. 中国皮革, 2008, 37(15): 8
- [39] 汤克勇, 刘京龙, 刘捷. 皮革耐干热收缩性能影响因素的研究[J]. 陕西科技大学学报, 2004, 22(3): 63-66

-
- [40] B.J.Alder, T.E.Wainright. Phase transition for a hard sphere system[J]. Journal of Chem Physics, 1957, 27: 1208-1209
- [41] A.W.Lees, Edwards. The computer study of transport proeess under extreme conditions [J]. J.Phys. C, 1957, C5: 1921-1929
- [42] 李春燕, 刘华, 刘波涛. 分子动力学模拟基本原理及研究进展[J]. 广州化工, 39 (4), 11-13
- [43] 杨萍, 孙益民. 分子动力学模拟方法及其应用[J]. 安徽师范大学学报(自然科学版), 2009, 1(32): 51-54
- [44] L.Verlet. Computer 'experiments' classical fluids: I. Thermodynamical properties of Lennard-Jones molecules [J]. Physical Review, 1967, 159(1): 98-103
- [45] Swope W C, Anderson H C, Berens P H, etal. A computer simulation method for the ealculation of equilibrium constants for the formation of physcial clusters of molecules: application to small water clusters [J]. Journal of Chemical Physics, 1982, 76: 637-649
- [46] R.W. Honeycutt. The potential calculation and some applications[J]. Methods in Computational Physics, 1970(9): 136-211
- [47] Beeman D. Some multistep methods for use in molecular dynamics calculations [J]. Journal of Computational Physics, 1976, 20: 130-139
- [48] 陈丽群. BCC-Fe 中刃型位错及其扭折的原子结构、电子结构和掺杂效应研究[D]: [硕士学位论文]. 北京: 钢铁研究总院, 2006
- [49] Rahman A. Correlations in themotion of atoms in liquid argon [J]. Physical Review A, 1964, 136: 405-411
- [50] Alder BJ, Walnwright TE. Phase transition for a hardsphere system [J]. The Journal of Chemical Physics, 1957, 27: 1208-1209
- [51] Rahman A. Correlations in the motion of atoms in liquid argon [J]. Physical Review A, 1964, 136: 405-411
- [52] Daw MS, Baskes MI. Embedded atom method: derivation and application to impurties, surfaces, and other defects in metals[J]. Phys Rev, 1984, B(29): 6443-6453
- [53] 文玉华, 朱如曾, 周富强等. 分子动力学模拟的主要技术[J]. 力学进展, 2003, 33(1): 66-68
- [54] 吕玲红. 烷烃在沸石分子筛中吸附分离的分子模拟研究[D]: [硕士学位论文]. 浙江: 浙江大学, 2005
- [55] H.J.C.Berendsen, J.P.M.Postma, W.F.V.Gunsteren. Moleeula: dynamies with coupling to an external bath [J]. J.Chem.Phy, 1984, 81: 3684-3690
- [56] S.A.Nose. Unified formulation of the constant ternperature moleeular dynamies methods [J]. J.Chem.Phy, 1984, 81: 5-11
- [57] 晋兴华. 分子模拟方法优化脂质纳米给药系统结构与性能的研究[D]: [硕士学位论文]

- 文]. 天津: 天津大学, 2009
- [58] 曹同成, 肖苏林, 赵国华. 分子动力学方法模拟热致蛋白质构象变异研究[J]. 第五届全国环境化学大会, 2009: 599
- [59] 蒲尚志, 张文华. 分子动力学模拟氨基酸序列对胶原蛋白热稳定性的影响[DB]. 中国科技论文在线(<http://www.paper.edu.cn>), 2012-1-12
- [60] Clark EDB. Refolding of recombinant proteins [J]. *Current Opinion in Biotechnology*, 1998, 9(2): 157-163
- [61] Liu Fu-Feng, Ji Luo, Dong Xiao-Yan. Effects of molecular volume and fractional polar surface area of osmolytes on the stability of chymotrypsin inhibitor 2 [J]. *Acta Phys-Chim.Sin.*, 2010, 26(10): 2813-2820
- [62] 蓝慰青. 猪皮胶原蛋白的提取及应用研究[D]: [硕士学位论文]. 上海: 上海水产大学, 2006
- [63] 赵帅, 李国英. 猪皮胶原的研究进展[J]. *皮革化工*, 2006, 23(2): 19
- [64] 姜冬宇. 蛋白质在聚乙烯界面吸附的分子动力学模拟[D]: [硕士学位论文]. 天津: 天津大学, 2008
- [65] 都文婕. 肽抑制剂稳定淀粉样蛋白构象的分子机理研究[D]: [硕士学位论文]. 天津: 天津大学, 2010

附录

附录 I 用于统计氢键的程序

```

function hbond_sum
clc
clear
hbond=load('1cgd_instrand.txt');
nh=zeros(length(hbond(:,1)),20);
for i=1:length(hbond(:,1))
    hl=find(hbond(i,:)~=0);
    index=hbond(i,hl);
    for j=1:length(index)/2
        nh(i,2*j-1)=index(j);
        nh(i,2*j)=index(j+length(index)/2);
    end
    nht(i)=length(index)/2;
end
nh
nht
hh=[nht',nh]
ii=62;
jj=3;
sij=0;
ht=[0,0];
for i=1:length(hbond(:,1))
    for j=1:10
        if nh(i,2*j-1)~=0&nh(i,2*j)~=0
            hti=find(nh(i,2*j-1)==ht(:,1));
            if length(hti)~=0
                htj=find(nh(i,2*j)==ht(hti,2));
                if length(htj)==0
                    ht=[ht;nh(i,2*j-1),nh(i,2*j)];
                end
            end
        else

```

```
        ht=[ht;nh(i,2*j-1),nh(i,2*j)];
    end
end
end
end

htn=[];
for k=1:length(ht)
    sij=0;
    for i=1:length(hbond(:,1))
        for j=1:10
            if nh(i,2*j-1)==ht(k,1)&nh(i,2*j)==ht(k,2)
                sij=sij+1;
            end
        end
    end
    htn(k,:)=[ht(k,1),ht(k,2),sij];
end
htn=htn(2:end,:)
length(htn)

end
```

附录 II 本人在攻读学位期间所发表的论文及获奖

- [1] 杜丛. 人造革合成革基布概述[J]. 魅力中国, 2010, 12: 395
- [2] 杜丛(第二作者). 利用精益生产解决服装企业中的浪费现象[J]. 武汉纺织大学学报, 2011, 24(2): 17-19
- [3] 杜丛(参与人). 内衣的人机工学研究与产品结构设计与产品结构. 湖北省教育厅科学技术研究计划重点项目. 2010
- [4] 杜丛(参与人). 纺织面料及服装数字化设计与制造成套技术研究. 纺织面料及服装数字化设计与制造成套技术研究. 2010
- [5] 杜丛(参与人). 高性能合成革服用材料研发及性能评价指标建立. 中国纺织工业协会科技指导性项目. 2007-2010