

第五章 蛋白质纯化、鉴定及
结构与功能分析-3

杜 芬

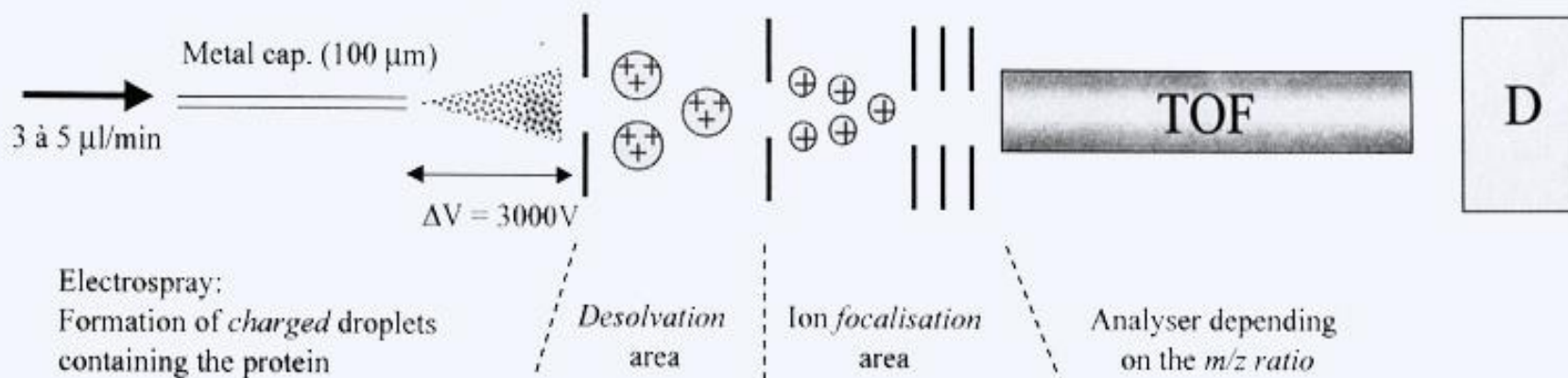
Email: fen.du@whu.edu.cn

4、生物质谱与蛋白质鉴定

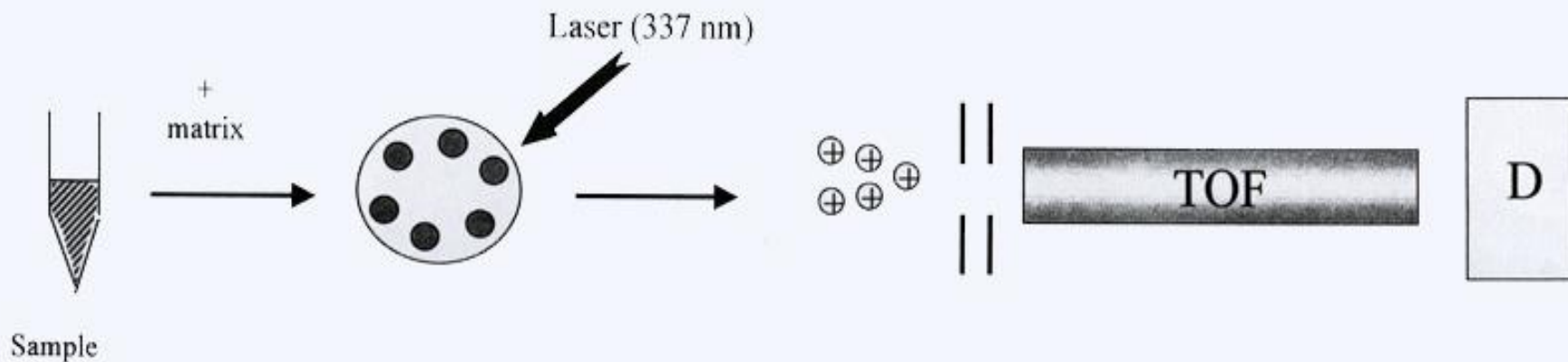
- ✿ 基质辅助的激光解吸/离子化(matrix assisted laser desorption/ ionization, MALDI), 另一种是电喷雾离子化(electrospray ionization, ESI)。这些技术能快速而极为准确地测定生物大分子的分子量;
- ✿ 再结合新的质谱分析技术(如飞行时间, Time-of-Flight, TOF), 便可在各种水平上研究蛋白质等生物大分子。

电喷雾离子化-飞行时间质谱

ESI-MS (electrospray)



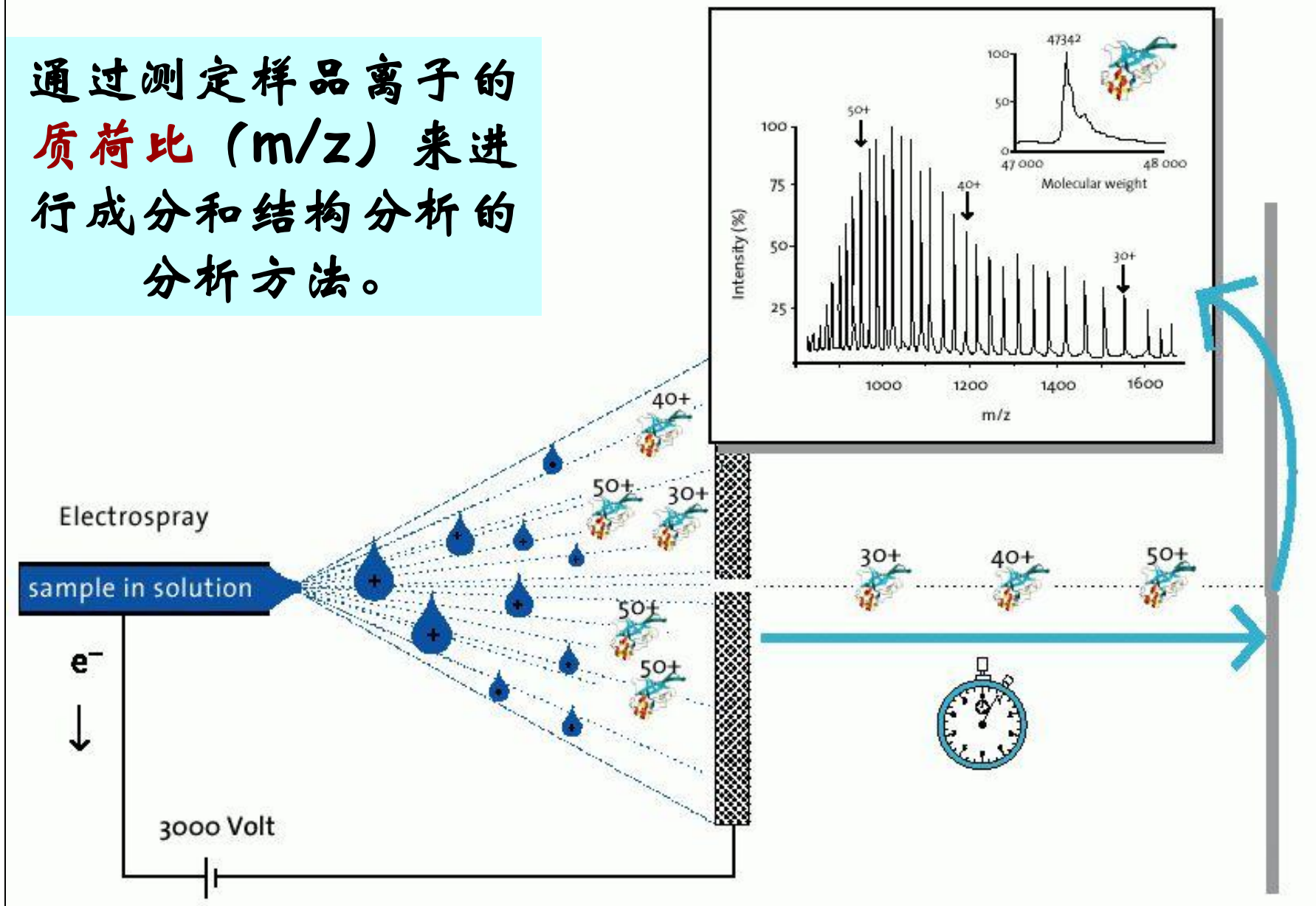
MALDI-MS (Matrix Assisted Laser Desorption Ionisation)



肌型肌酸激酶

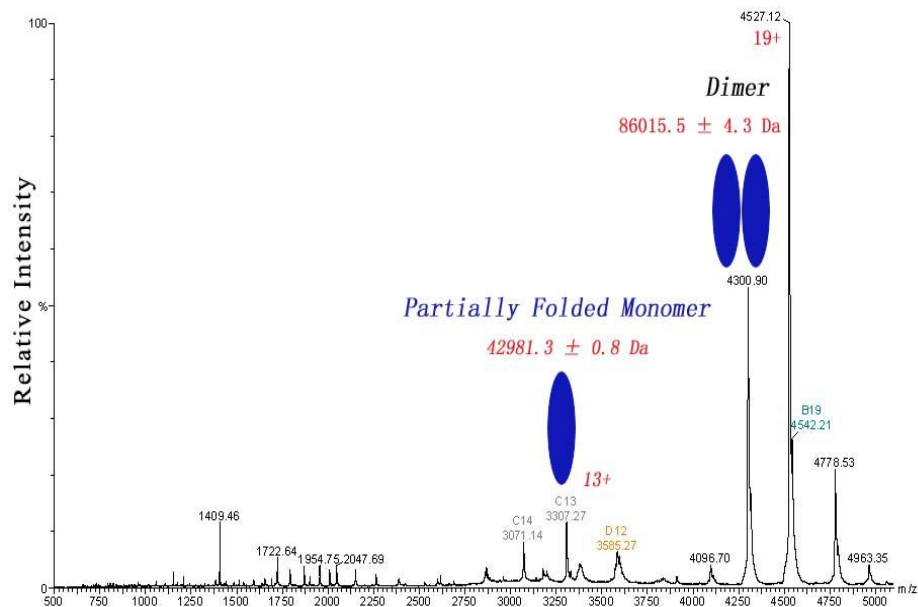


通过测定样品离子的
质荷比 (m/z) 来进行
成分和结构分析的
分析方法。



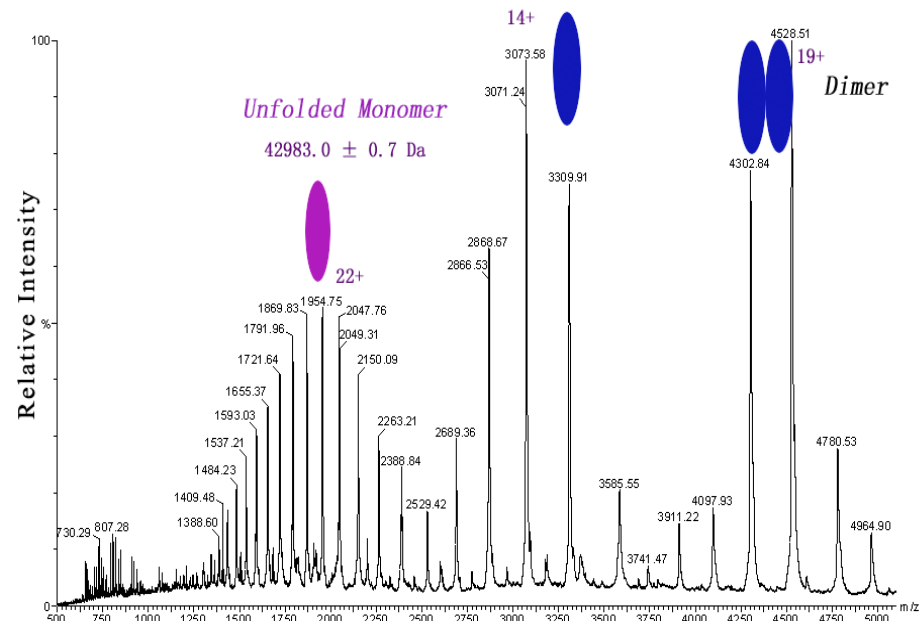
电喷雾电离质谱分析原理

pH = 6.7

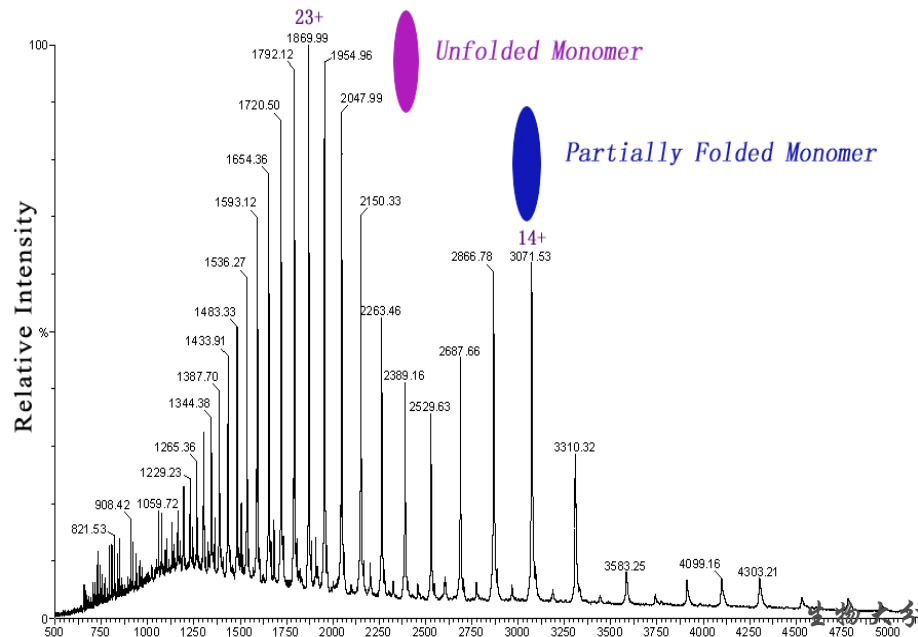


pH = 5.0

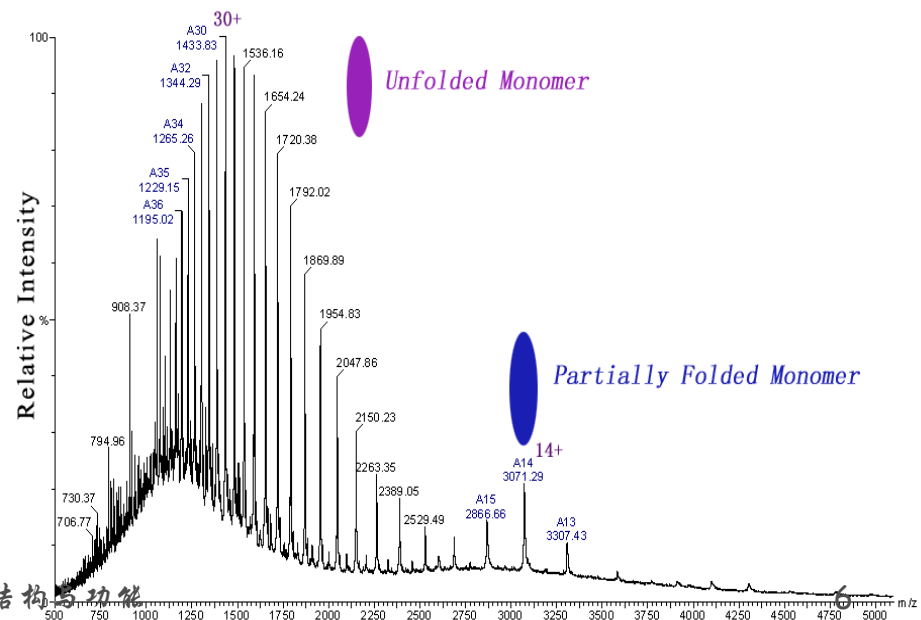
Partially Folded Monomer

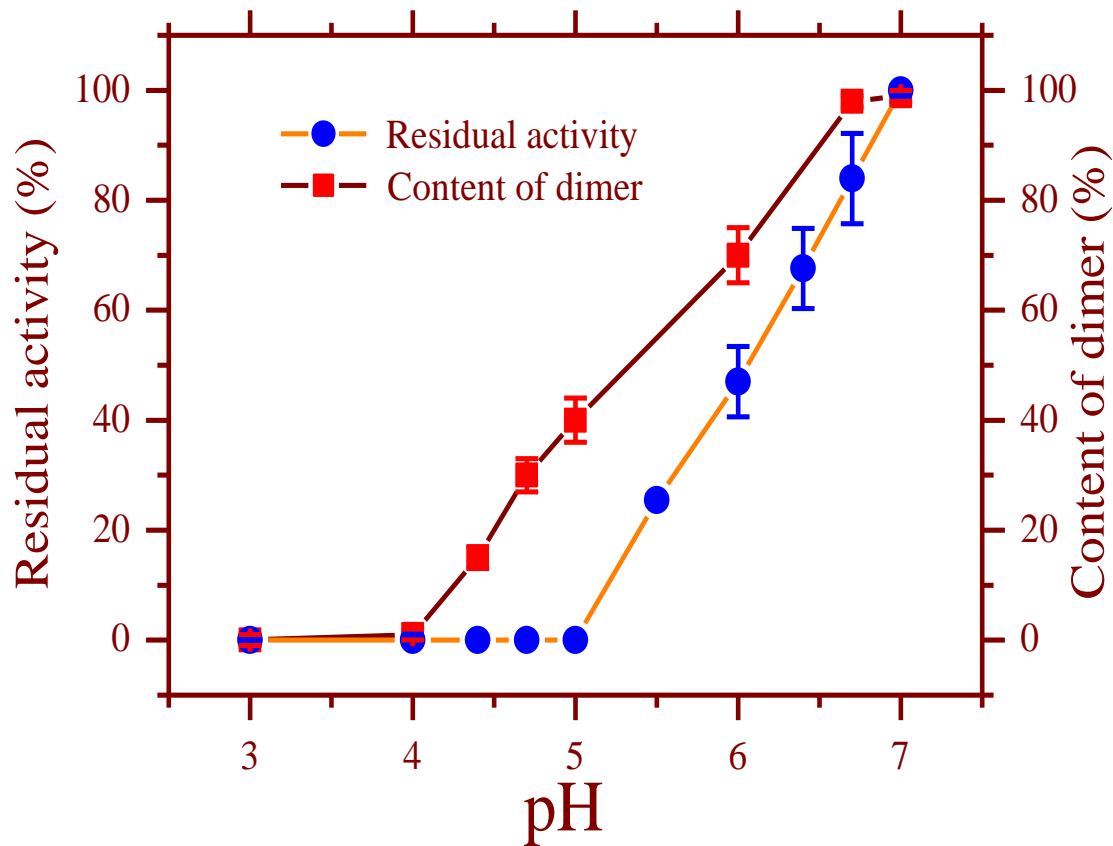


pH = 4.0



pH = 3.0





● 结果表明，在酸诱导蛋白质变性过程中，MM-CK活力丧失先于四级结构的改变，结合质谱图证据可说明该酶的活性部位比酶分子整体更容易遭到破坏 (Tsou, 1993)。

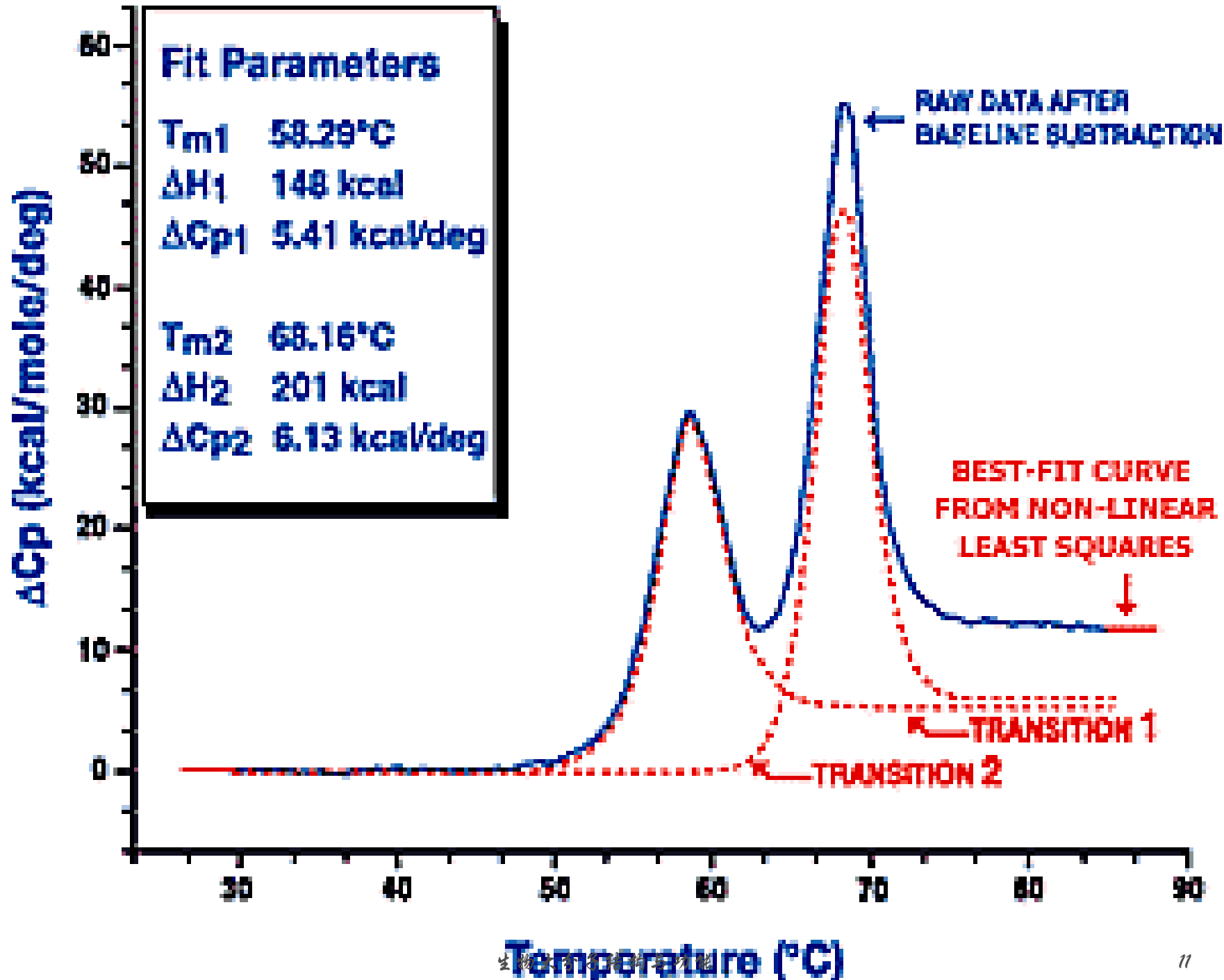
5、差示扫描量热法

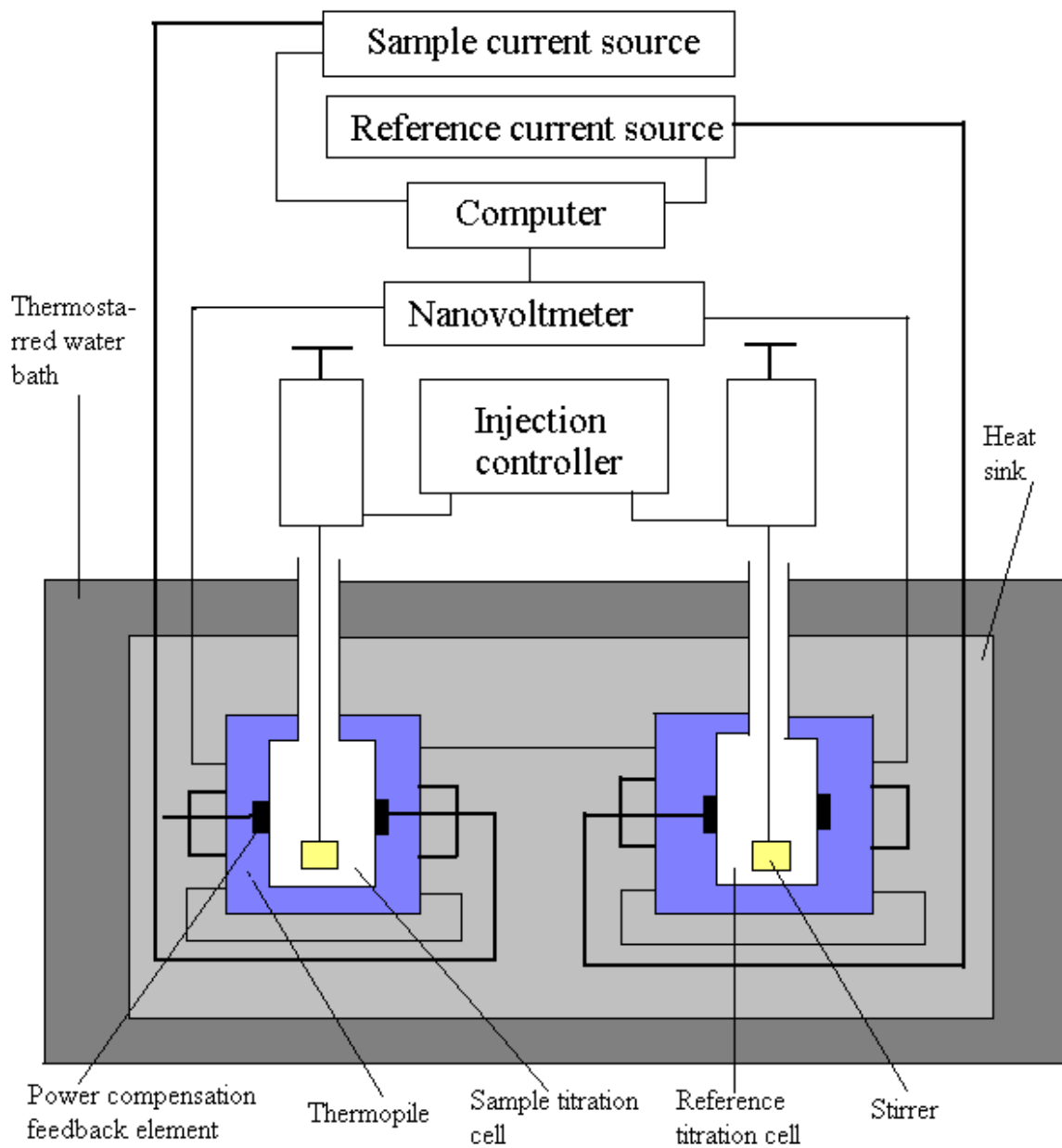
- ❁ 差示扫描量热法是20世纪60年代以后研制出来的一种热分析方法。
- ❁ 所谓差示扫描量热技术，是指在样品和参比同时程序升温或降温且保持两者温度相等的条件下，测量流入或流出样品和参比物的热量差与温度关系的一种技术。

- ❁ 这种热量差是与生物大分子如蛋白质、核酸等的构象能量的不同有关的。
- ❁ 通过测量不同温度下转型的发生及转型产生的热变化，可以获得生物大分子结构方面的信息。
- ❁ 通过程序降温可以获得有关变性过程的可逆性的相关信息。
- ❁ 根据测量方法的不同，可分为两种类型：热流型DSC和功率补偿型DSC。

MicroCal DSC





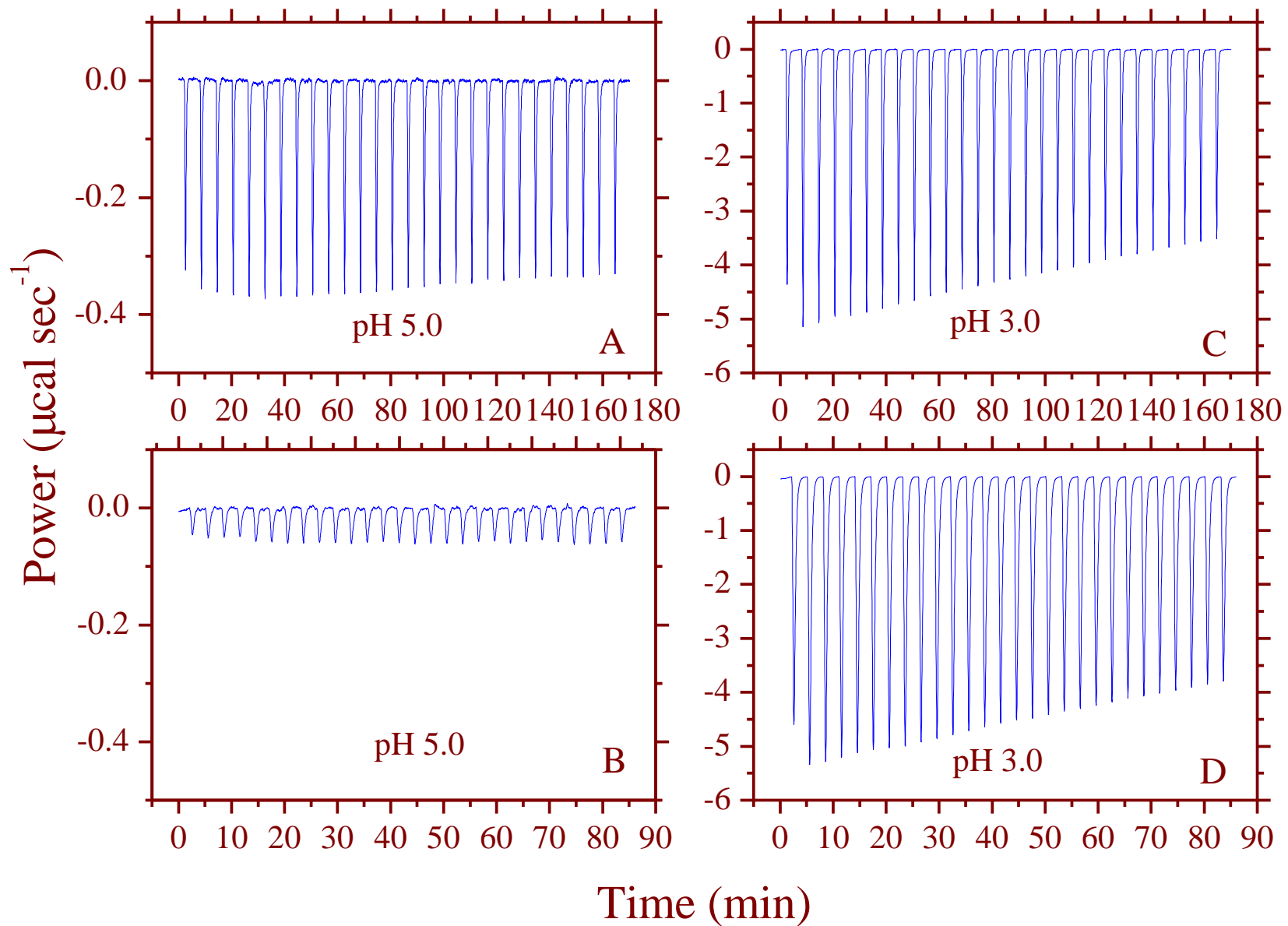


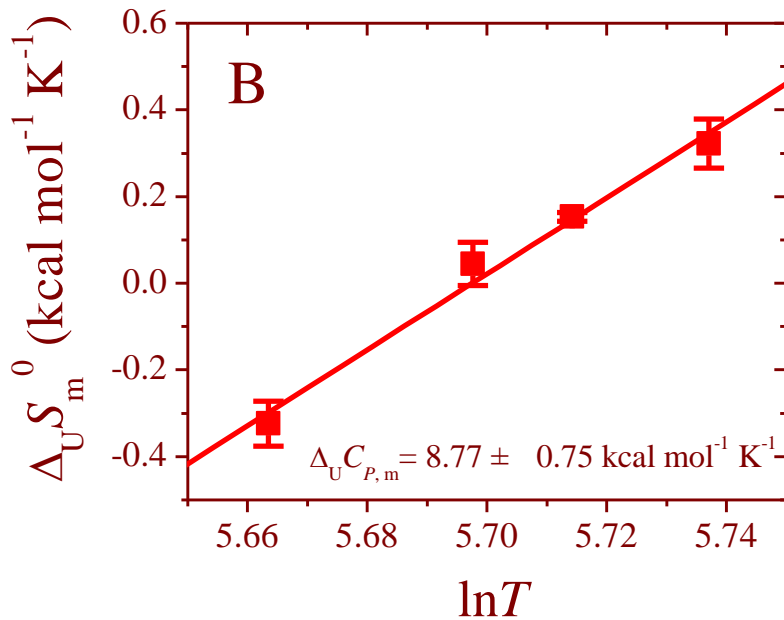
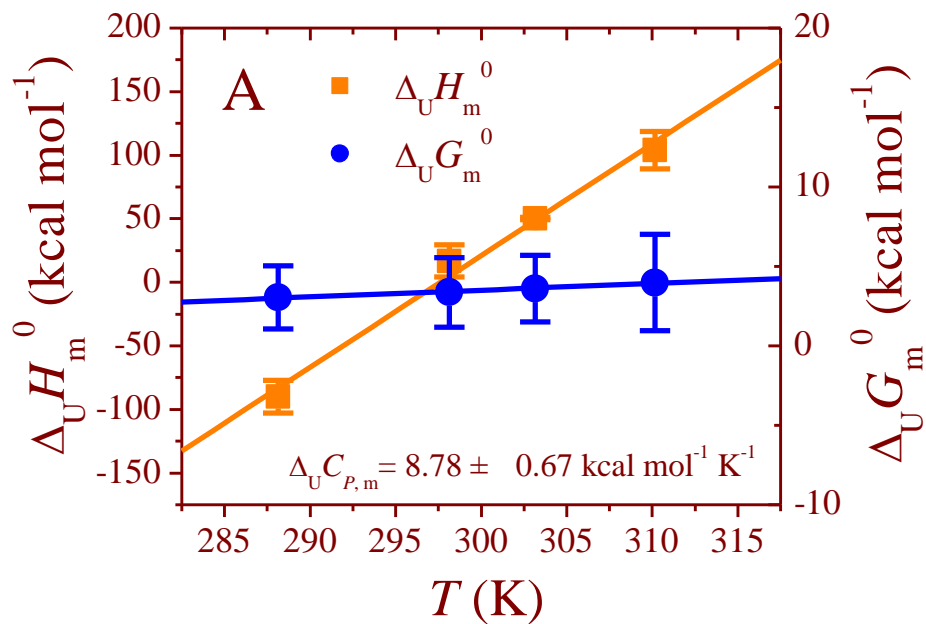
6、等温滴定量热法 (Isothermal Titration Calorimetry, ITC)



Measurements

- Affinity (K_f)
relative strength of molecular interactions
- Enthalpy (ΔH)
heat absorbed or lost
- Entropy (ΔS)
relative disorder
- Heat Capacity (C_p)
types of surfaces interacting





该蛋白质去折叠的摩尔热容变化在所研究的范围内是温度不依赖的。

✿ ITC做蛋白质相互作用的图

7、荧光光谱

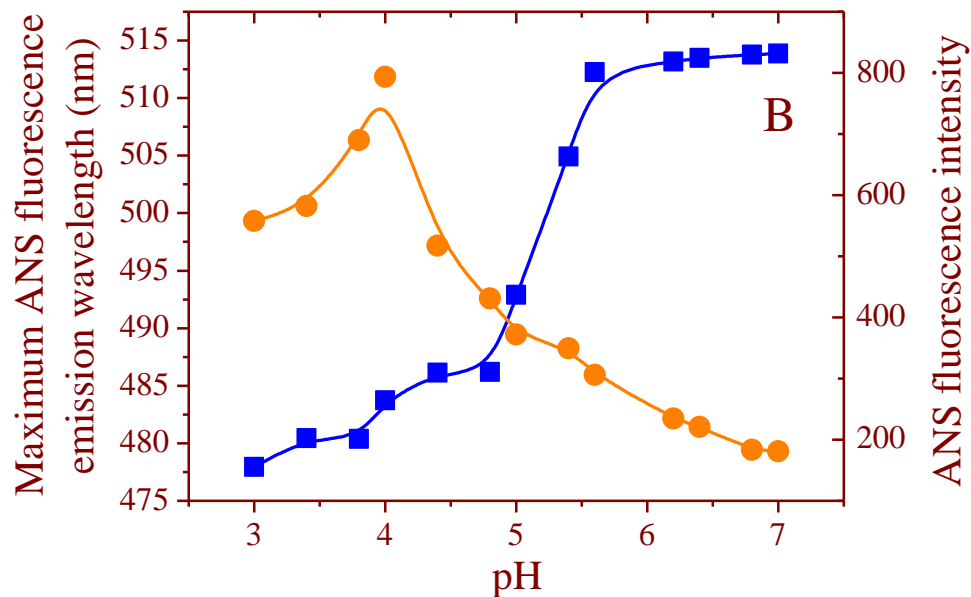
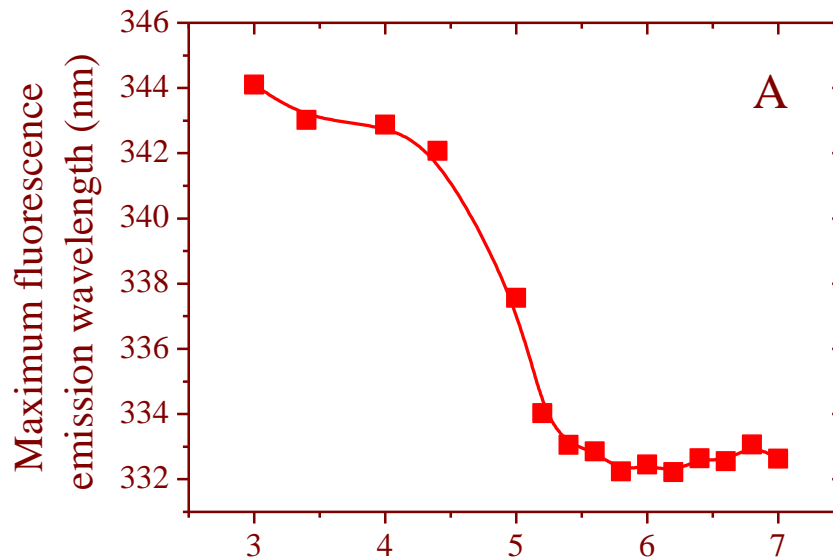
内源荧光

激发波长: 295nm

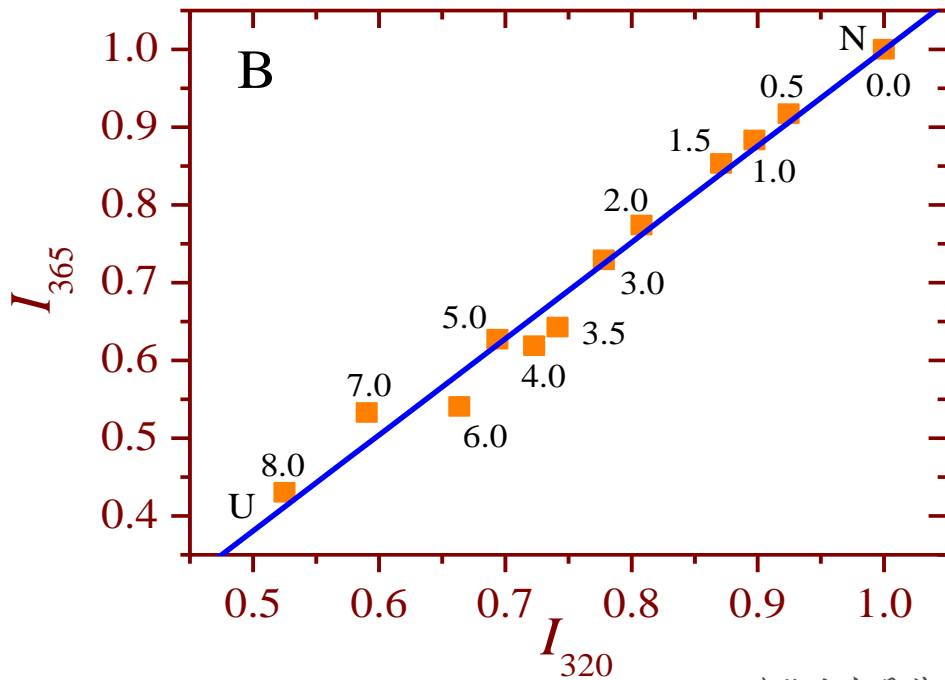
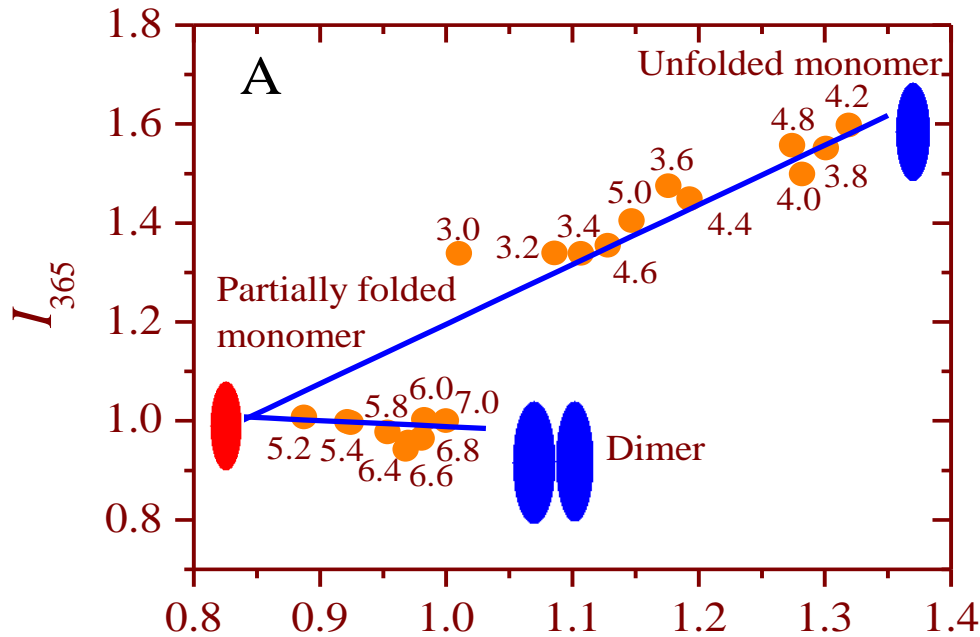
色氨酸

外源荧光法

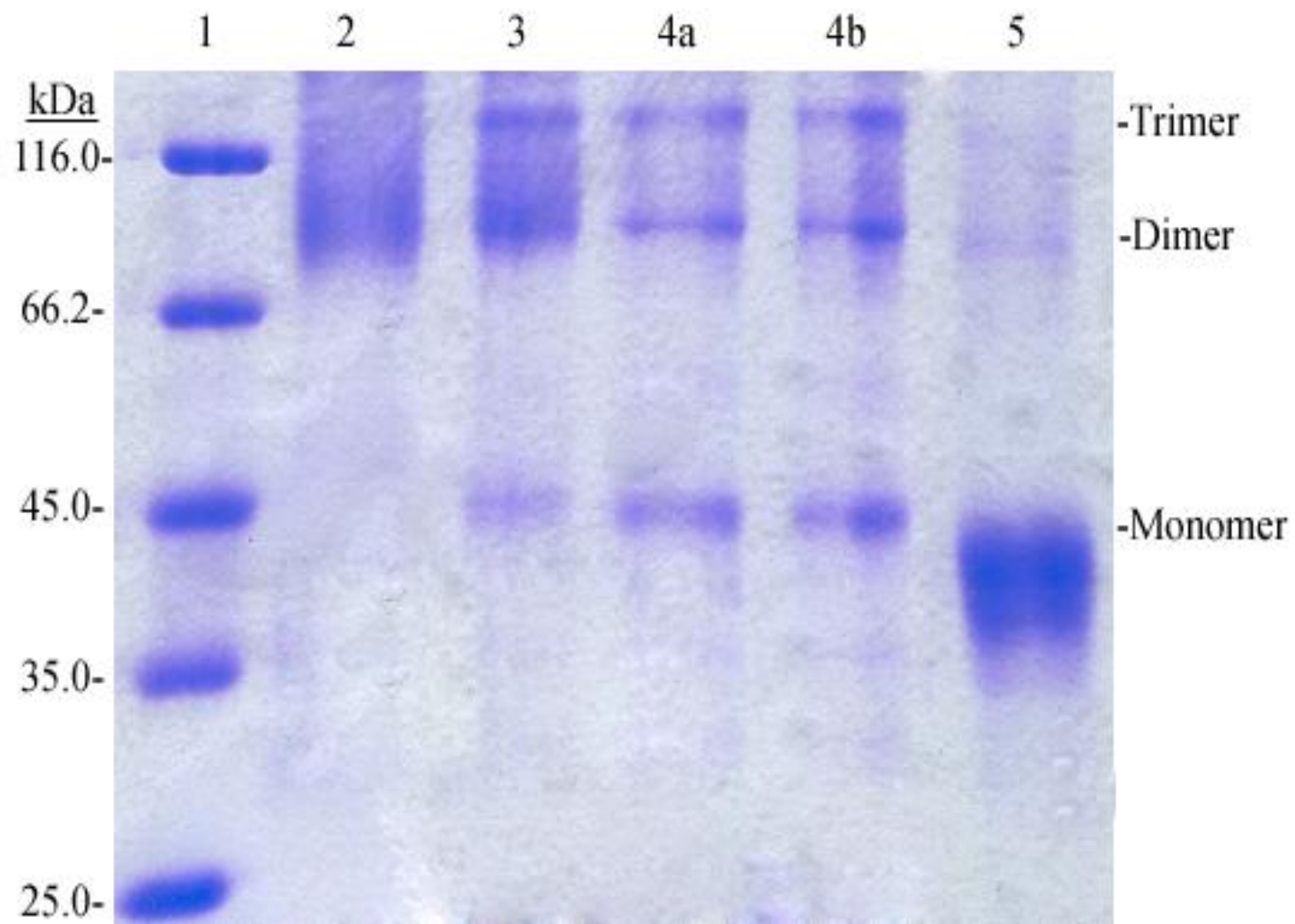
激发波长: 390 nm



相图法



✿ 酸诱导 MM-CK 去折叠过程符合“三态模型”：天然态(二体蛋白) → 中间体(部分折叠的单体，结合ESI-TOF MS的结果) → 去折叠态(去折叠的单体)。



8、戊二醛交联电泳法

J Biol Chem. 2003;278:30098-105.

傅立叶变换红外光谱法 (FTIS)

(1) FTIS研究蛋白质的二级结构，主要是对其红外光谱中酰胺I谱带进行分析。酰胺I谱带为 α -helix、 β -strand、 β -turn和coil等不同结构振动峰的加和带，彼此重叠，在1260-1700 cm^{-1} 范围内通常为一个不易分辨的宽带谱。目前常用去卷积、微积分等数学方法，使加和带中处于不同波数的二级结构的各个吸收峰得以辨认。最后经过谱带拟合，获得各个吸收峰的信息。

(2) FITS对蛋白质的纯度的要求：>95%的纯度。

生物信息学预测蛋白质高级结构

(1) 同源建模法 (modeling) :

(2) 折叠识别法 (threading) :

(3) 从头预测法 (*de novo*) :

❁ 物理性质预测:

<http://expaxy.hcuge.ch/ch2d/pi-tool.html>

Peptidemass

<http://expaxy.hcuge.ch/sprot/peptide-mass.html>

TGREASE <ftp://ftp.virginia.edu/pub/fasta/> SAPS

http://ulrec3.unil.ch/software/SAPS_form.html

Expasy 中 ProtParam

(<http://web.expasy.org/protparam>)

❁ 基于组成的蛋白质识别预测

AACompIdent

<http://expaxy.hcuge.ch/ch2d/aacompi.html>

AACompSim

<http://expaxy.hcuge.ch/ch2d/aacsim.html>

PROPSEARCH <http://www.embl->

[heidelberg.de/prs.html](http://www.embl-heidelberg.de/prs.html)

❁ 二级结构和折叠类预测

nnpredict

<http://www.cmpharm.ucsf.edu/~nomi/nnpredict>

Predictprotein <http://www.embl->

[heidelberg.de/predictprotein/SOPMA](http://www.embl-heidelberg.de/predictprotein/SOPMA)

<http://www.ibcp.fr/predict.html> SSPRED

<http://www.embl->

[heidelberg.de/sspred/ssprd_info.html](http://www.embl-heidelberg.de/sspred/ssprd_info.html)

❁ 特殊结构或结构预测

COILS

http://ulrec3.unil.ch/software/COILS_form.htm

[MacStripe](http://ulrec3.unil.ch/software/COILS_form.html)

<http://www.wi.mit.edu/matsudaira/macstripe.html>

[ml](http://www.wi.mit.edu/matsudaira/macstripe.html)

❁ 蛋白质三维建模: SWISS-MODEL

(<http://swissmodel.expasy.org/workspace>)

- ✿ 疏水性分析: ExPASy的ProtScale程序
(<http://www.expasy.org/cgi-bin/protscale.pl>), 软件
bioedit, dnamana等
- ✿ 跨膜区分析
 - ✗ 存在的跨膜螺旋Tmbase 数据库 (<http://www.isrec.isb-sib.ch/ftp-server/tmbase>)
 - ✗ 蛋白质跨膜区预测: Tmpred
(http://www.ch.embnet.org/software/TMPRED_form.html)
- ✿ 前导肽与蛋白质定位
<http://genome.cbs.dtu.dk/services/signalP-2.0> 蛋白质细胞
定位: TargetP 1.1
(<http://www.cbs.dtu.dk/services/TargetP>)
- ✿ 卷曲螺旋分析coiled-coil
<http://www.york.ac.uk/depts/biol/units/coils/coilcoil.html>
COILS
http://www.ch.embnet.org/software/COILS_form.html Ep
itopeInfo <http://epitope-informatics.com/Links.htm> 蛋白
质二级结构预测: SWISS-MODEL
(<http://swissmodel.expasy.org/workspace>)

- ❁ 蛋白序列同源性分析：
NCBI/BLAST(www.ncbi.nlm.gov/blast), 选择程序BLASTP就可网上分析, 华盛顿大学的BLAST软件 (dove.embl-heidelberg.de/blast2)
- ❁ Motif数据库-PROSITE:
www.expasy.org/prosite
蛋白质序列的(profile)分析
www.isrec.isb-sib.ch/software/PFSCAN_form.html
- ❁ InterProScan综合分析网站
www.ebi.ac.uk/interpro/scan.html
- ❁ 蛋白质的结构功能域分析
简单模块构架搜索工具 (simple modular architecture research tool, SMART) 一个较好的蛋白质结构功能域的数据, 可用于蛋白质结构功能域的分析, 所得到的结构域同时提供相关的资源的链接<http://smart.embl-heidelberg.de/>

蛋白质结构预测

✿ PDB数据库

蛋白质基本立体结构数据库(PDB, www.rcsb.org)其中有大量工具用于查看PDB数据库中的结构, 如rasmol, 可用于显于出蛋白质的空间结构, 下载地址: www.umass.edu/microbio/rasmol)

✿ PDBFinder 数据库

是在PDB、DSSP、HSSP基础上建立的二级库, 它包含PDB序列, 作者, R因子, 分辨率、二级结构等, 这些些信息随着PDB库每次发布新版, PDBFinder在EBI自动生成, 网址为“www.sander.embl-heidelberg.de/pdbfinder/ [ftp://swift.embl-heidelberg.de/pdbfinder.](ftp://swift.embl-heidelberg.de/pdbfinder)

✿ NRL-3D数据库

是所有已知结构蛋白质的数据库, 可用于查询蛋白序列时行相似性分析以确定其结构www.gdb.org/Dan/protein/nrl3d.html

ISSD数据库

蛋白质序列数据库, 其每个条目包含一个基因的编码序列, 同相应的氨基酸序列对比, 并给出相应的多肽链结构数据。

www.protein.bio.msu.su/issd

HSSP数据库

是根据同源性导出的蛋白质二级结构数据库, 每一条PDB项目都有一个对应的HSSP文件, www.sander.embl-heidelberg.de/hssp

❁ 蛋白质结构分类数据库

对已知蛋白质三维结构进行手工分类得到的数据库，位于剑桥的站点也提供BLAST检索服务
<http://scop.mrc-lmb.cam.ac.uk/scop/>

MMDB蛋白质分子模型数据库是ENTREZ检索工具所使用的三维结构数据库，以ASN格式反纳的PDB中的结构和序列数据。NCBI同时提供一个配套的三维结构显示程序的Cn3D, www.ncbi.nlm.nih.gov/Structure/

Dali/FSSP数据库基于PDB数据库中现有的蛋白质三维结构，用自动结构对比程序Dali比较而形成的折叠单元和家庭分类库。 www.embl-ebi.ac.uk/dali

❁ 蛋白质二级结构预测
PHD程序提供了从二级结构到折叠方面分析的多种资源
www.embl-heidelberg.de/predictprotein/predictprotein.html, 也可通过
[email:predictprotein@embl-heidelberg.de](mailto:predictprotein@embl-heidelberg.de) 进行数据分析。

❁ BCM Search Launcher
<http://searchlauncher.bcm.tmc.edu/seq-search/>
蛋白质序列二级结构预测综合站点, 从此出发, 输入蛋白质序列, 可以根据需要, 使用各种在线预测工具, 包括
Coils、nnPredict、PSSP/SSP、PSSP/NNSSP、SAPS、TMpred、SOUSI、Paircoil、Protein
Hydrophilicity/Hydrophobicity Search、SOPM, 使用十分方便。

❁ 蛋白质分类数据库 (ProtoMap)
是对SWISS-PROT数据库中的全部蛋白质由计算机自动进行层次分类, 把相关者聚集分级所得到的数据库。
www.proteinmap.cs.huji.ac.il

✿ DAS

<<http://www.biokemi.su.se/~server/DAS/>>;

蛋白跨膜预测服务器、输入蛋白序列，预测跨膜区域。

✿ TopPred 2

<<http://www.biokemi.su.se/~server/toppred2/>>;

斯德哥尔摩大学理论化学蛋白预测服务器提供的膜蛋白拓扑学预测在线工具

✿ SOSUI

<<http://azusa.proteome.bio.tuat.ac.jp/sosui/>>;

东京农业科技大学提供的膜蛋白分类和二级结构预测在线工具。

✿ PSIPred - MEMSAT2

<<http://globin.bio.warwick.ac.uk/psipred/>>;

✿ HMMTOP

<<http://www.enzim.hu/hmmtop/>>;

预测蛋白序列的跨膜螺旋与拓扑结构服务器

✿ TMpred
<http://www.ch.embnet.org/software/TMPRED_form.html>;
预测蛋白序列跨膜区

TMHMM
<<http://www.cbs.dtu.dk/services/TMHMM-1.0/>>;
预测蛋白的跨膜螺旋

The PredictProtein server
<<http://www.embl-heidelberg.de/predictprotein/>>;
提供蛋白数据库查询，预测蛋白各种结构的服务

SMART
<<http://smart.embl-heidelberg.de/>>;
提供蛋白序列，在结构域数据库中查询，显示出其结构域及跨膜区等

SPLIT
<<http://pref.etfos.hr/split/>>;
膜蛋白二级结构预测服务器

PRED-TMR
<<http://o2.db.uoa.gr/PRED-TMR/>>;
提供基于SwissProt数据库统计分析的预测蛋白跨膜片段的服务

CoPreThi
<<http://o2.db.uoa.gr/CoPreTHi/Main.html>>;
基于INTERNET的JAVA程序，预测蛋白的跨膜区



TMAP

http://www.embl-heidelberg.de/tmap/tmap_info.html;

提供预测蛋白跨膜片段的服务

multalin

<http://www.toulouse.inra.fr/multalin.html>;

蛋白序列对照服务器，比较几条蛋白序列的结构。

Protein Sequence Analysis

<http://bioweb.pasteur.fr/seqanal/protein/intro-uk.html>;

巴斯德研究所提供的常用蛋白序列分析在线工具，绝对精选。

PSA

<http://bmerc-www.bu.edu/psa/>;

Protein Sequence Analysis 服务器为美国波士顿大学生物分子工程研究中心 (the BioMolecular Engineering Research Center) 开发，提交氨基酸序列，预测二级结构及折叠区域。

PRS

<http://www.embl-heidelberg.de/prs.html>;

EMBL提供的以未知蛋白序列的氨基酸组成而非氨基酸序列顺序进行蛋白家族及各种特性预测的服务器，并将结果通过EMAIL发给您。

AAA

<http://www.embl-heidelberg.de/aaa.html>;

EMBL氨基酸分析服务器，与上一个服务类似，以氨基酸残基组成为蛋白分析基础数据，结合蛋白数据库进行分析。

www.ncbi.nlm.nih.gov



Nucleotide: insulin



NM_000207: Homo sapiens insulin (INS), mRNA



agccctccag gacaggctgc atcagaagag gccatcaage agatcactgt cttctgcc
tgccctgtg gatgcgcctc ctgccctgc tggcgctgct ggccctctgg ggacctgacc
cagccgcage ctttgtgaac caacacctgt gcggctcaca cctggtggaa gctctctacc
tagtgtgcgg ggaacgaggc ttcttctaca cacccaagac ccgccgggag gcagaggacc
tgcaggtggg gcaggtggag ctggggcgggg gccctggtgc aggcagcctg cagccctgg
ccctggaggg gtccctgcag aagcgtggca ttgtggaaca atgctgtacc agcatctgct
cccttacca gctggagaac tactgcaact agacgcagcc cgcaggcage cccacaccg
ccgcctcctg caccgagaga gatggaataa agcccttgaa ccagcaaaa



MALWMRLLPLLALLALWGPDPAAAFVNQHLCGSHLVEAL
YLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQP
LALEGLSLQKRGIVEQCCTSICSLYQLENYCN



www.expasy.org



Swiss-Prot :Protein knowledgebase



ExPASy Proteomics tools



ProtParam tool



Number of amino acids: 110

Molecular weight: 11980.9

Theoretical pI: 5.22



Amino acid composition:

Ala (A) 10 9.1%; Arg (R) 5 4.5% ; Asn (N) 3 2.7% ; Asp (D) 2 1.8% ; Cys (C) 6 5.5%
Gln (Q) 7 6.4% ; Glu (E) 8 7.3% ; Gly (G) 12 10.9%; His (H) 2 1.8% ; Ile (I) 2 1.8% Leu
(L) 20 18.2% ; Lys (K) 2 1.8% ; Met (M) 2 1.8% ; Phe (F) 3 2.7%; Pro (P) 6 5.5% Ser (S)
5 4.5% ; Thr (T) 3 2.7% ; Trp (W) 2 1.8% ; Tyr (Y) 4 3.6% ; Val (V) 6 5.5% ;

Total number of negatively charged residues (Asp + Glu): 10

Total number of positively charged residues (Arg + Lys): 7

- ✿ www.expasy.ch/enzyme An enzyme database.
- ✿ users.rcn.com/jkimball.ma.ultranet/BiologyPages/A/Apoptosis.html Highly textual description of apoptosis.
- ✿ www.biology.arizona.edu/biochemistry/problem_sets/energy_enzymes_catalysis/energy_enzymes_catalysis.html An extensive description of enzyme catalysis