



生物大分子三维结构显示技术讲义

PyMOL 使用入门

耿存亮

gengcunliang@gmail.com

2012 年 10 月 09 日

1. PyMOL 简介

PyMOL是一款生物大分子三维结构显示软件，其中“Py”是指此软件使用Python 语言编写，“MOL”是指Molecule。

PyMOL官网是 <http://www.PyMOL.org/>，发展历史和软件更新动态可在此查询。

PyMOL的学习网站是 http://www.PyMOLwiki.org/index.php/Main_Page，若想知道用好PyMOL，此网站是必上网站，其实这一个也就够了。

2. PyMOL 入门

2.1 模式显示及颜色显示

PyMOL既可以鼠标操作也可以命令操作，但是命令操作可以完成许多鼠标难以完成的任务。下面就以实例来认识一下PyMOL。

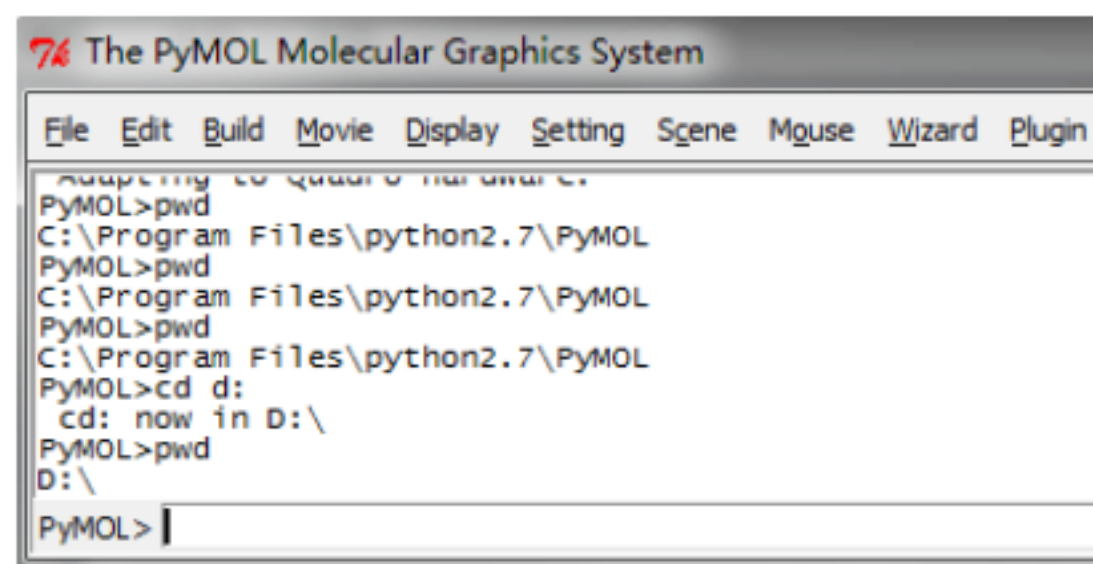
打开PyMOL软件后，首先要特别注意的是当前工作路径。
在命令框输入命令并回车

```
pwd
```

即可显示当前工作路径，默认路径是PyMOL的安装路径。一般不把文件保存在安装路径下，所以需修改当前工作路径，而且路径不能有汉字，
比如改为D盘，输入命令并回车

```
cd d:
```

再用pwd命令查看一下当前工作路径。如下图所示：



pwd: print working direction

cd: change direction

命令很方便很简单很神奇吧 O(_)O~

其次要注意的是保证鼠标是三键式的，滚轮可用作中键。如果像苹果机一样

只有一个按键或没有中键的话，还是赶紧换个鼠标吧。

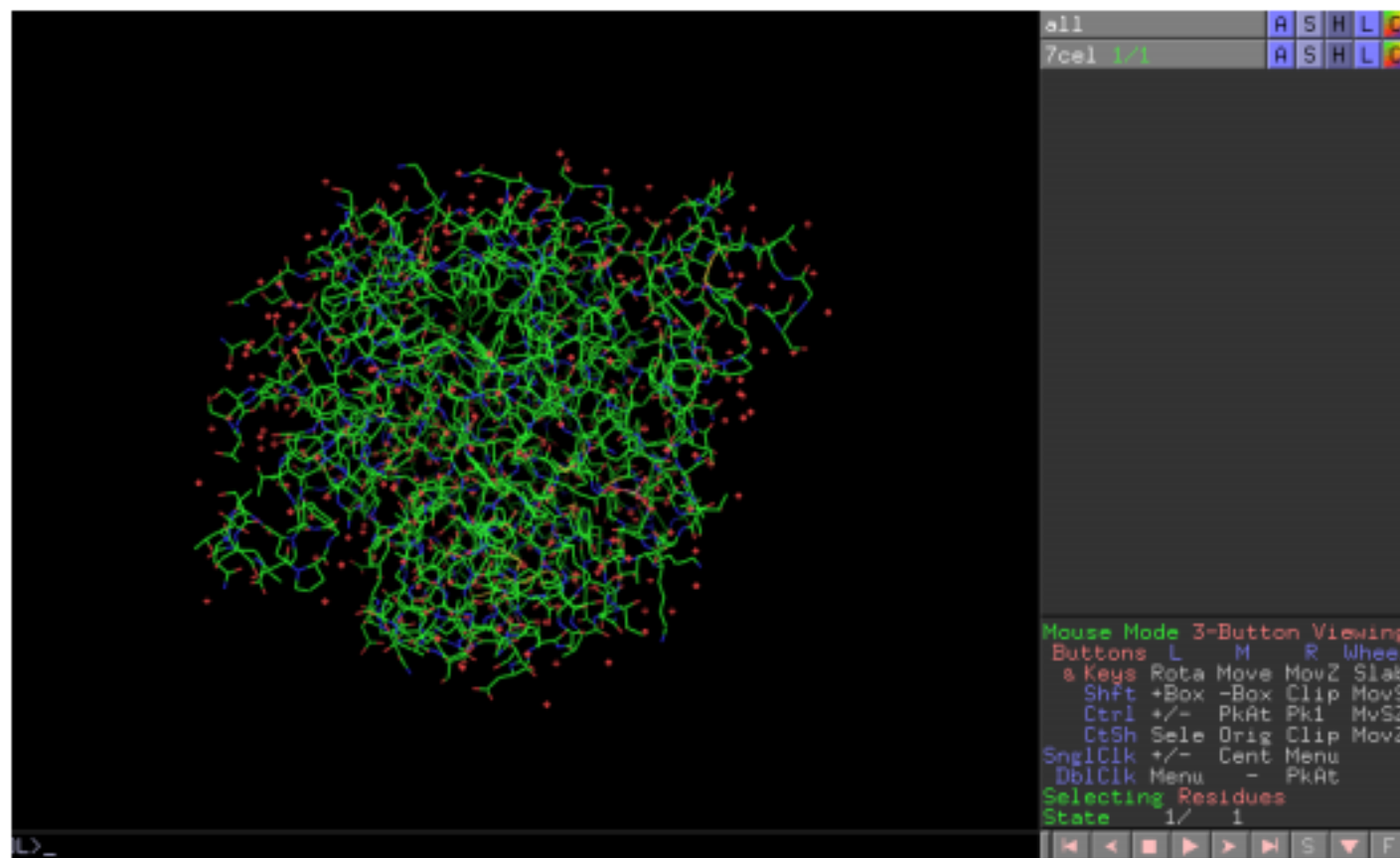
好了，现在下载一个 PDB 文件，如何下载呢？

当然可以去 PDB 网站 <http://www.rcsb.org/pdb/home/home.do> 下载，但是打开网页多麻烦啊，如果能用 PyMOL 直接下载该多好啊，那就试一试 fetch 命令吧！

下载纤维素外切酶 CBHI 和纤维素糖链的复合物晶体结构，PDB 号是 7cel，输入命令

```
fetch 7cel
```

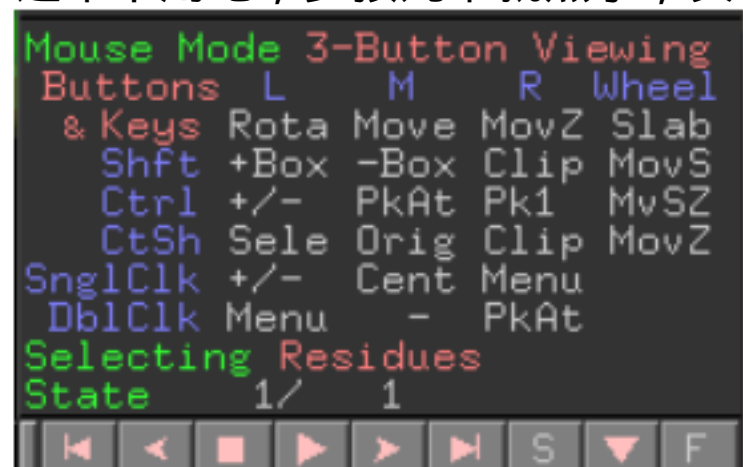
稍等片刻，就会下载完毕并显示如下：



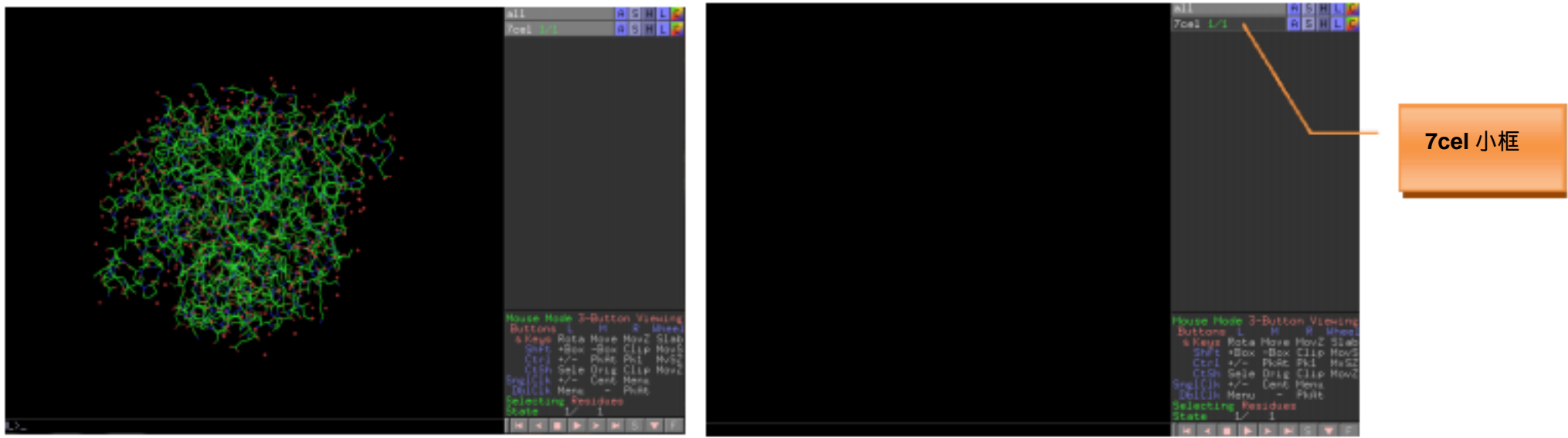
刚才说到三键式鼠标，那么三个键都有什么用呢？

按住左键滑动会旋转结构（ rotate ），按住中键滑动会移动结构（ move ），按住右键滑动会缩放结构（ move zoom ）。

这个不用记，多按几下就熟了，实在忘了在右下角有提示的，如下图

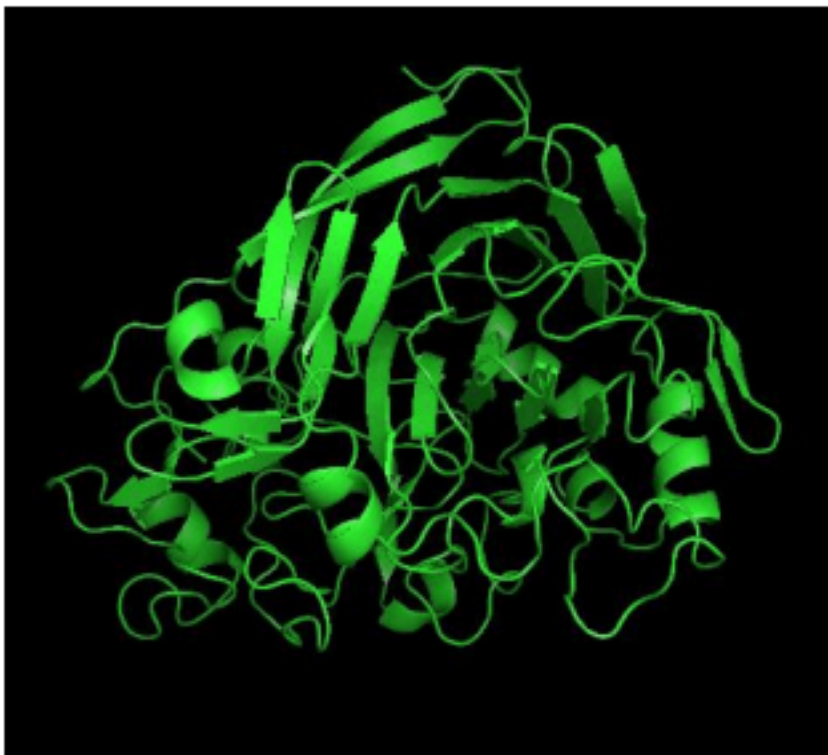


下载打开 7cel 的 pdb 文件后，在 all 小框下面会出现 7cel 小框，后面还跟着几个按键 ASHLC，这几个按键后面会一点点介绍。先左键点击一下 7cel 小框，会发现结构消失了，被点击的小框也变暗了，这就是隐藏功能，再点击一下就会恢复，如下图所示

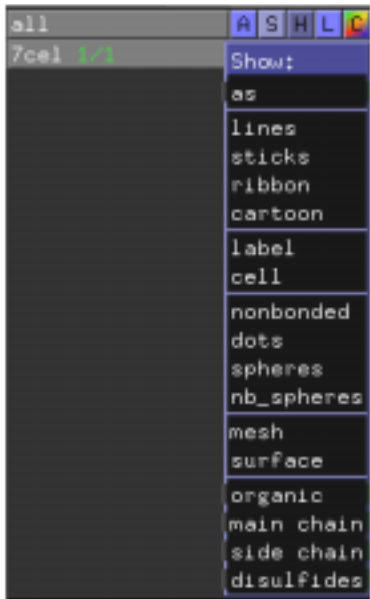


看到恢复后的结构你一定感到一团糟吧，这都神马呀！
上过王禄山老师的课后，应该清楚 PDB文件就是具有一定格式的文本文档，里面记录着每个原子的三维坐标值，不信的话可用记事本打开 PDB文件看一看。PyMOL所谓的结构显示就是读取每个原子的三维坐标，然后用一个点（球）来显示出来，原子之间的化学键用线表示。当然还可以用其他模式的显示，比如大家很熟悉的螺旋飘带模型，在 PyMOL中叫 cartoon 模式，输入命令并回车

as cartoon



看着熟悉的 螺旋和 折叠，是不是感觉清爽多了？
刚才的操作也可用鼠标完成，如下图所示



点击 7cel 小框中的 S 按键，然后再点击 cartoon 或者 lines、sticks、ribbon 等等。假如点击了 lines，你会发现一团糟的结构又回来了，而 cartoon 模式没有

消失。是的，这就是 show 和 as 的不同，show 是指显示一种或多种模式，而 as 是指只显示为一种模式。as 该点击哪里呢？仔细看一下上图就明了了。同时你也可能明白了 S 按键是 Show 的意思。

命令显示多种模式

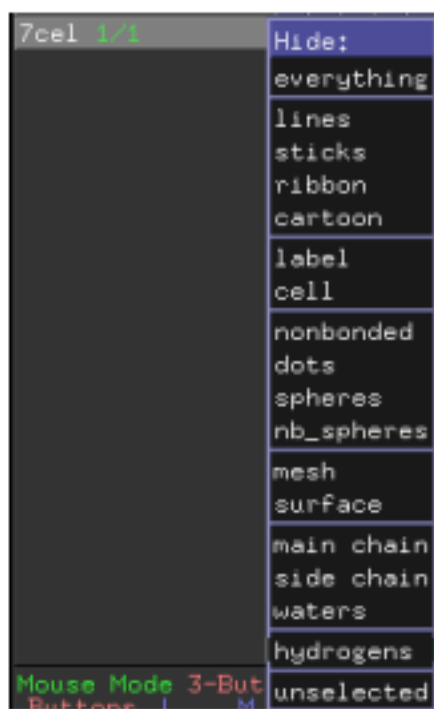
show lines

show sticks

show ribbon

怎么隐藏不想显示的模式呢？H 按键就能办到，H 是 Hide 的意思。点击 H 中的相应模式，就能使其隐藏。当然了命令也能办到，比如隐藏 lines 模式

hide lines



好了，现在输入

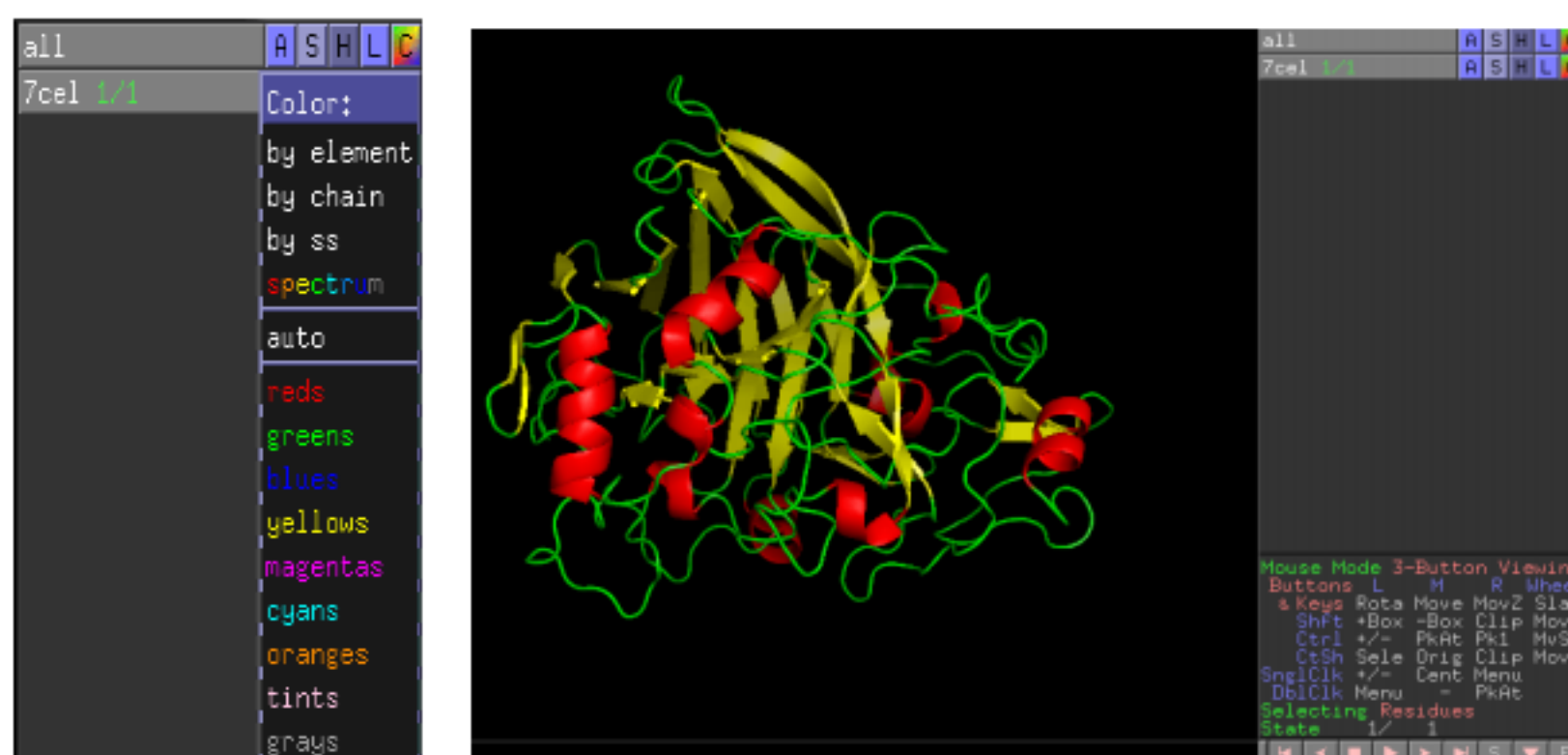
as cartoon

只显示 cartoon 模式，蛋白的三级结构显示地很清楚，二级结构中的螺旋、折叠和无规则卷曲也很清楚地显示了。

咦？怎么蛋白是绿色的？难道蛋白真的是绿色的吗？蛋白到底是什么颜色我不清楚，这里显示的颜色是软件设置的默认颜色，既然是软件设置，那么当然可以改成其他颜色。这就要用到 C 按键，C 是 Color 的意思。

点击 C 按键你会看到下图所示的工具框，其中 by element 指按元素类型着色，by chain 按肽链着色，by ss 按二级结构着色（secondary structure），spectrum 是指渐变色，还有 red、green 等等各种颜色。

点击一下 by ss，你会发现螺旋、折叠和无规则卷曲显示成了不同的颜色，这样显示整个结构是不是更清楚了。



那么怎么用命令实现着色呢？比如整个蛋白显示绿色，命令是

```
color green
```

对 螺旋、 折叠和无规则卷曲着不同的颜色怎么实现呢？

比如 螺旋 (helix) 显示红色， 折叠 (sheet) 显示绿色，无规则卷曲 (loop) 显示蓝色，命令是

```
color red,ss h
```

```
color green,ss s
```

```
color blue,ss l+
```

想必你猜出来 ss h, ss s的意思了，其实就是选择二级结构 -helix 或 -sheet ,那么 ss l+ ” (这里单引号和双引号都行)为什么还要写 + ” 呢？ 因为二级结构的分类不仅仅只有这三种， PyMOL为了简化，就把不是 -helix 和 -sheet 的结构全部归为 loop 和其他无规则结构了，所以这里得用 l+ ” 表示。

文献上的结构图一般都是白色背景，怎么设置背景颜色呢？ 命令是

```
bg_color white
```

也可以改成其他颜色， 但是发文章默认都是白色背景， 在电脑上看一般是黑色背景，这样设置是有道理的， 喜欢探究的同学可以查一下印刷和屏幕的颜色显示原理。

学到此，先复习一下，温故而知新
模式显示或隐藏

```
as cartoon ( lines、 sticks、 ribbon、 surface、 spheres、 mesh 等等 )
```

```
show cartoon
```

```
hide cartoon
```

颜色设置

```
color red
```

```
color blue,ss s
```

背景颜色设置

```
bg_color white
```

2.2 选择命令

如果只想选一个原子或一个残基或一段结构，然后突出显示，这该怎么办呢？这个问题很关键，这涉及到 PyMOL 中很重要的选择命令。

比如选择 7cel 中的一个催化残基 212 位的谷氨酸 GLU，命令如下

```
select geng,resi 212
```

Select 是选择命令，后面紧跟一个名字 geng，就是给所选择的残基起个名字，这个随便起，然后 resi 212 是指残基号为 212 的残基，resi 是 residue id 的缩写。

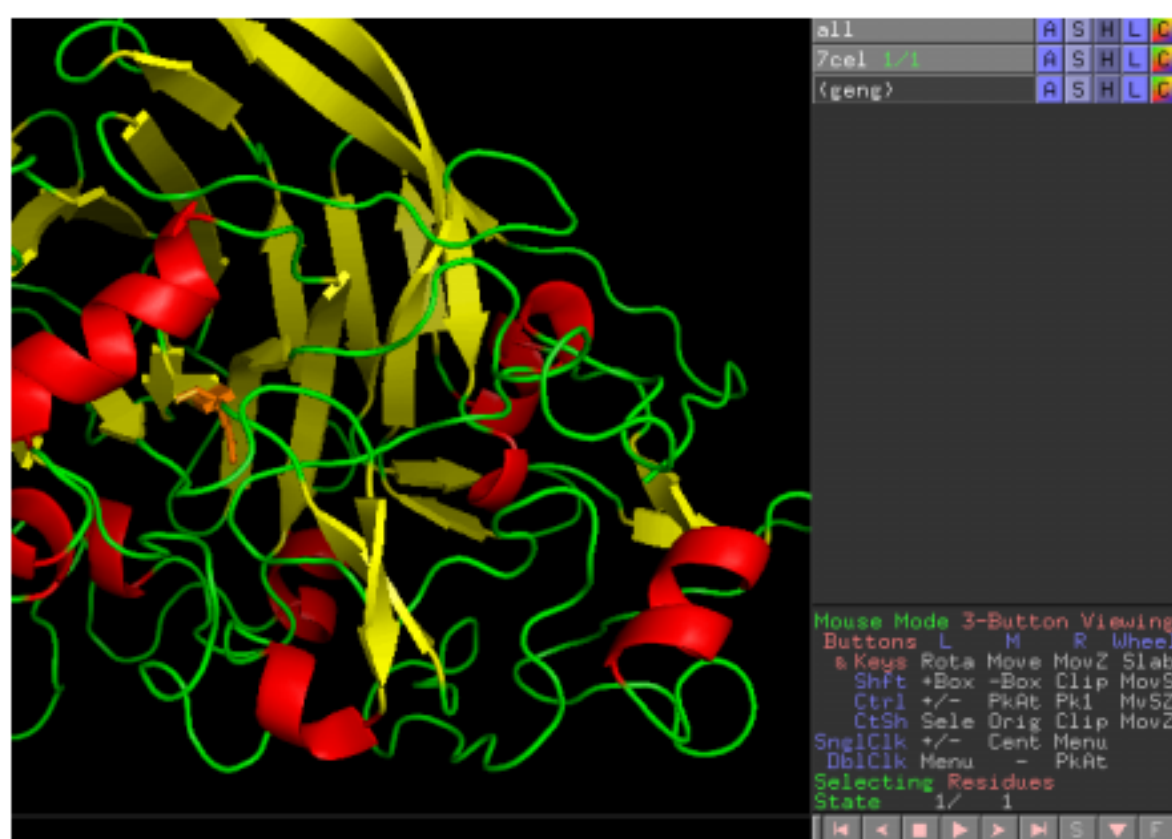
输入命令并回车后，结构上会显示出一些粉红色的点，7cel 框下面会出现一个 geng 框。但是你还是没看到 212GLU 的样子，怎么办呢？

点击 geng 框后 S 按键中的 sticks，212GLU 残基就会显示出 sticks 模式，再点击 C 按键中的 oranges 改为桔色，如下图。这一切也可用命令实现

```
show sticks,geng
```

```
color orange,geng
```

看出来给所选残基起名字的好处了吧，这样就可以在命令中指定对谁操作，如果什么都不指定那就是对所有原子进行操作，也就是上一节学的对整个蛋白改模式改颜色。



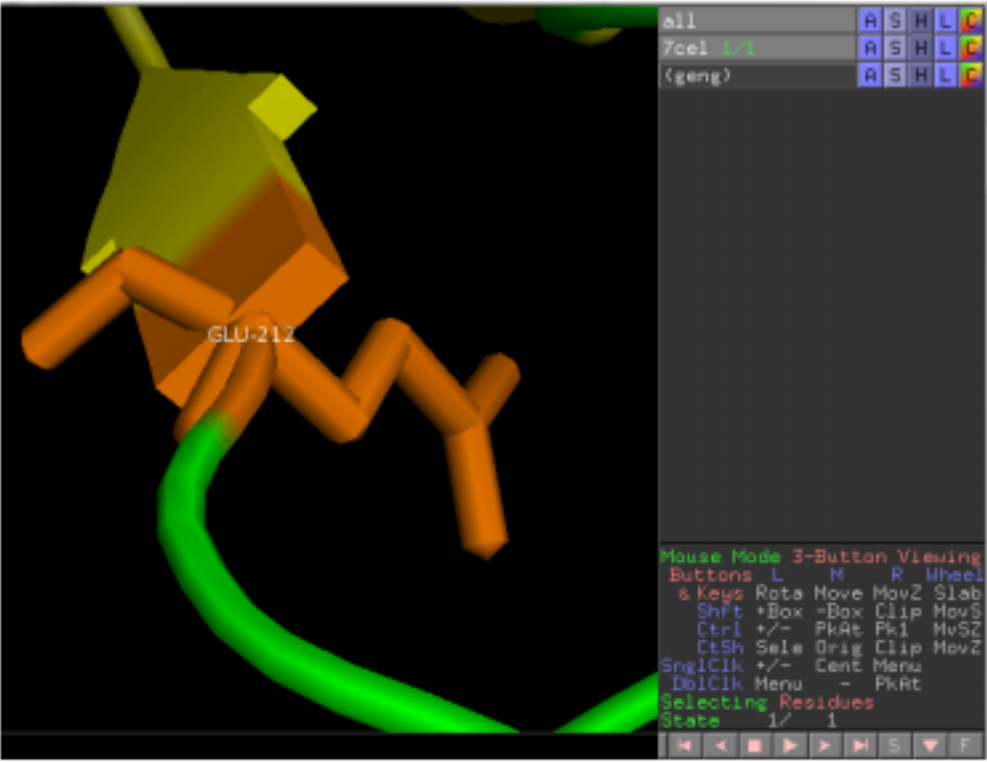
刚才选择了 212GLU，如何突出显示并标记出来呢？

点击 geng 小框 L 按键中的 residues，这样就对 212GLU 加上了标签，L 就是 Label 的意思。

然后输入命令

```
zoom geng
```

这样就突出显示 212GLU残基了，如图



点击上图右下角的 S 按键，S 指 sequence，点击后会显示出蛋白序列。刚才选择残基是用命令实现的，也可以用鼠标左键在蛋白结构或蛋白序列上点击，点击中的残基就会被选中并起名为 sele，然后使用 A 按键中的 rename selection 修改名称即可，你会发现被点中的残基都会显示出粉红色的小点。

可是到目前为止，所选择的最小单位都是残基，如果只想选择一个原子或者一条肽链，怎么做？鼠标操作的话，首先要修改选择的单位，仔细看一下右下角是不是有 Selecting Residues，点击 Residues，看看都有什么。如果想选择单个原子，那么点击到显示 Atoms 的时候停下来，如图



然后再在结构上点击某个原子，就选中它了。此时不能在序列上点击了，因为序列只有残基名没有原子名。

举一反三，你也可以选择一个分子、一条链、一个 C 原子等等。

用命令怎么实现自由选择呢？

除了能够根据残基号 resi 进行选择，还应该其他的吧？是的，这东西叫属性选择符，列表如下

属性选择符	缩略形式	标识符和举例
symbol	e.	Chemical-symbol-list 单字母或双字母的化学元素符号 PyMOL>select polar,symbol o+n
name	n.	Atom-name-list 蛋白和核酸中至多 4 字母的原子符号 PyMOL>select carbons,name ca+cb+cg
resn	r.	Residue-name-list

		3 字母的氨基酸符号 PyMOL>select aas,resn asp+glu+asn+gln 或至多 2 字母的核苷酸符号 PyMOL>select bases,resn a+g
resi	i.	Residue-identifier-list 至多 4 位数的残基号 PyMOL>select boy,resi 1+10+100+1000 Residue-identifier-range PyMOL>select boy,resi 1-10
alt	alt	Alternate-conformation-identifier-list 单字母 PyMOL>select altconf,alt a+ ' '
chain	c.	Chain-identifier-list 单字母或有时是数字 PyMOL>select 007,chain a
segi	s.	Segment-identifier-list 至多 4 字母 PyMOL>select ligand,segi lig
flag	f.	Flag-number 从 0 到 31 的单整数 PyMOL>select f1,flag 0
numeric_type	nt.	Type-number 单整数 PyMOL>select f1,nt. 5
text_type	tt.	Type-string 至多 4 字母 PyMOL>select subset,text_type HA+HC
id	id	External-index-number 单整数 PyMOL>select idno,id 23
index	idx.	Internal-index-number 单整数 PyMOL>select intid,index 11
ss	ss	Secondary-structure-type 单字母 PyMOL>select allstrs,ss h+s+l+ ' '

学到此，先复习一下，温故而知新嘛
选择操作

select geng,resi 212	
select geng,resn glu	选择所有 glu 残基并起名为 geng
select geng,name c+n+o	选择所有 c、n、o 原子并起名为 geng
select geng,chain a	选择 a 链并起名为 geng

```
''
```

突出显示

```
zoom resi 212
```

突出显示 212 号残基

2.3 布尔算符和标签命令

中学时学过布尔算符 `and/or/not`，在 PyMOL 的选择命令中布尔算符非常有用。比如选择 7cel 中 212GLU 残基的 C 原子，怎么选？

`select geng, resi 212` 和 `select geng, name ca` 两个命令显然不行，因为它们分别对 212 残基的所有原子以及蛋白的所有 C 原子进行了选择，而不只是 212GLU 的 C 原子。

可以这样做

```
select geng, resi 212 and name ca
```

那么如何选择 212 和 214 号残基的所有原子呢？

```
select geng, resi 212 or resi 214
```

选择除 200-400 号残基外的所有残基

```
select geng, not resi 200-400
```

```
color white, geng
```

 显示一下被选中的残基

选择 212 和 214 号残基的非 C 原子

```
select geng, resi 212+214 and not name ca
```

```
hide all
```

```
show sticks, geng
```

```
label geng, name
```

看一看是不是没有显示 C 原子的标签。

上一节做标签是使用鼠标操作的，这里用 `label` 命令完成，想隐藏标签直接输入

```
label
```

回车即可。

如果想知道如何用命令做其他标签，比如残基名残基号，甚至自己起的名字，可以输入

```
help label
```

然后你能看到下图所示的帮助信息了。 其他命令的帮助信息也可通过同样的方式获得，比如 `help color`、`help select`、`help show` 等等。

```
74 The PyMOL Molecular Graphics System
File Edit Build Movie Display Setting Scene Mouse Wizard Plugin

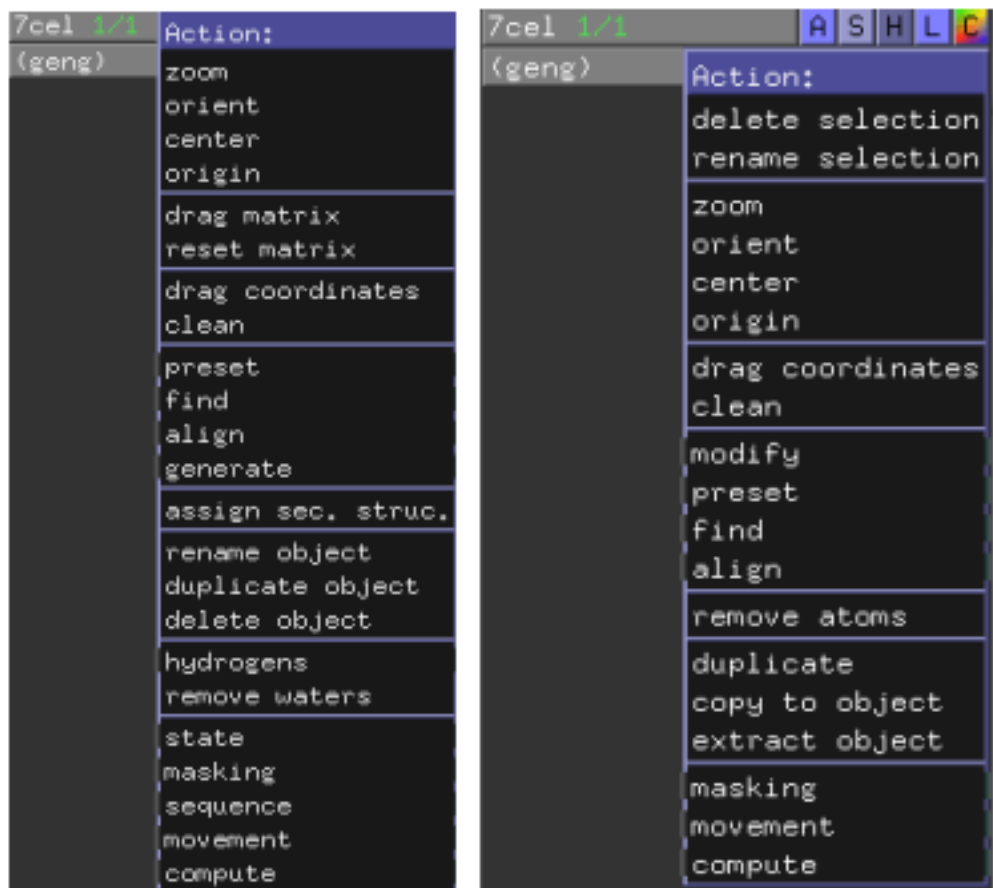
PyMOL>help label
DESCRIPTION
    "label" labels one or more atoms in a selection by evaluating an
    Python expression referencing properties for each atom.
USAGE
    label [ selection [, expression ] ]
ARGUMENTS
    selection = string: a selection-expression
    expression = string: a Python expression that can be converted to a string
EXAMPLES
    label chain A, chain
    label name ca,"%s-%s" % (resn,resi)
    label resi 200,"%1.3f" % partial_charge
NOTES
    The symbols defined in the label name space for each atom are:
        name, resi, resn, resv, chain, segi, model, alt, q, b, type,
        index, rank, ID, ss, vdw, elec_radius, label, elem, geom,
        flags, color, cartoon, valence, formal_charge, partial_charge,
        numeric_type, text_type, stereo
    All strings in the expression must be explicitly quoted.
    This operation typically takes several seconds per thousand atoms
    labelled.
    To clear labels, simply omit the expression or set it to ''.

PyMOL>
```

到此为止，基本操作就算学完了，已经学过了如何选择如何显示各种模式如何着色如何做标签，一副精美而有科学意义的图片是由基本命令组合使用完成的，这一点只能多用多练。

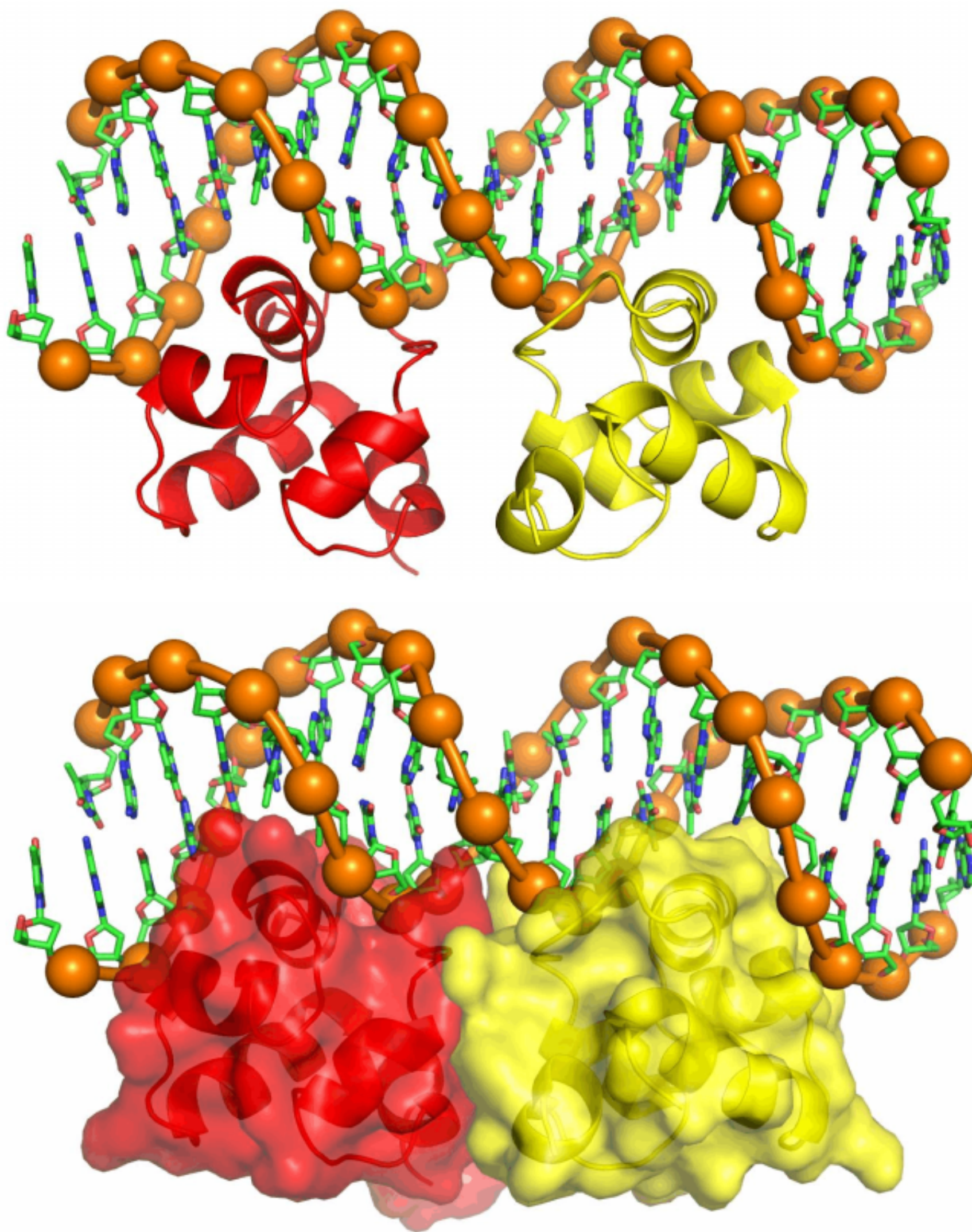
2.4 Object和 selection

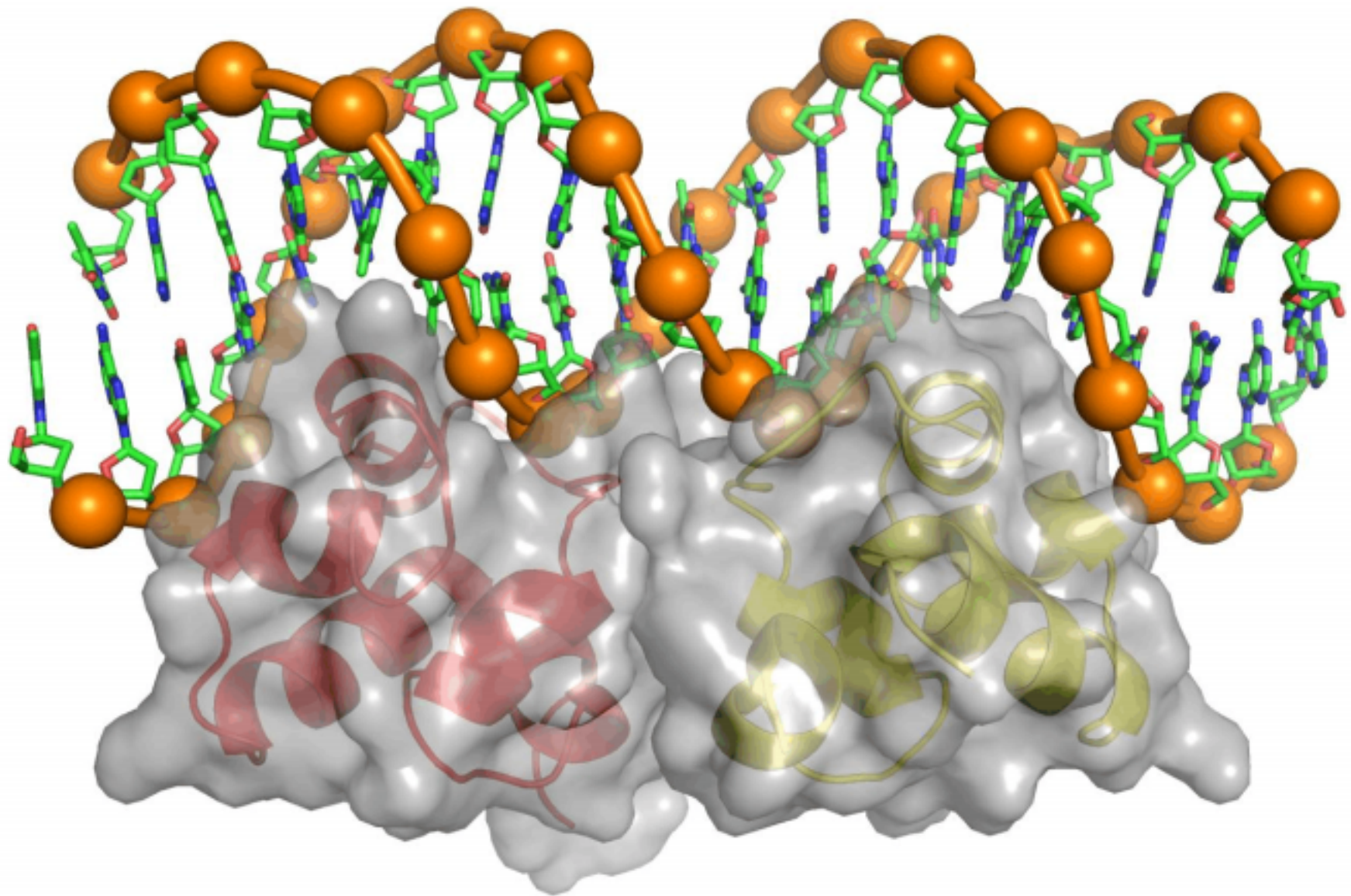
PyMOL中还有一点非常重要但是解释起来挺麻烦，我试着阐述一下。最初下载 7cel 生成一个 7cel 小框，后来做选择时也生成了一个 geng 小框，这两者有区别吗？可以点击一下 A 按键，看看两者显示的工具框是否不一样，如下图所示



PyMOL中管 7cel 叫 Object，而所做的选择叫 selection。两者什么区别呢？Object 是实实在在的物体或东西，而 selection 是一种虚指。Object 删除后，selection 也就消失了；而 selection 删除后被选择的原子不会从蛋白中消失，对 object 没有任何影响，所删除的只是一个指代名称而已。

管他 object 还是 selection，这对显示有什么影响呢？很有影响， 在一些操作中，这两个概念区分不开，常常达不到需要的效果。下面咱们就做几张精美图片，组合运用一下基本操作，练练手。





2.5 实例操练

上面三张图展示的是噬菌体的一种阻遏蛋白和 DNA 结合的晶体结构，这个阻遏蛋白结合到 DNA 后可阻止 DNA 的表达，PDB 号是 3cro。

下面一步步解释操作过程

删除 7cel 结构或所有结构

```
delete 7cel
```

```
delete all
```

下载 PDB 3cro

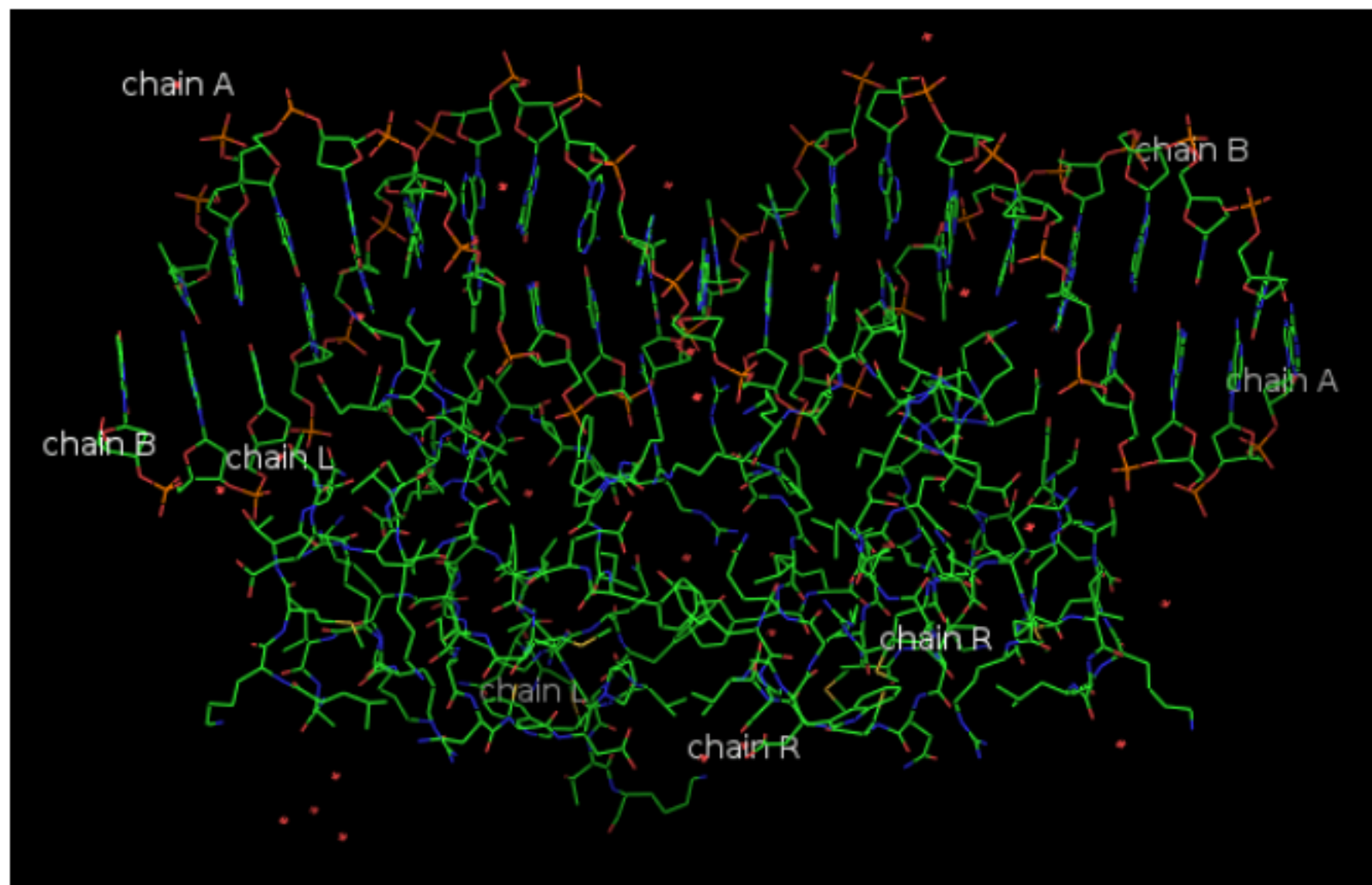
```
fetch 3cro
```

调整蛋白的姿势

```
orient
```

点击 S 按键查看序列中有几条链，或者点击 L 按键中的 Chains 使其显示出来，由图可知 DNA 有 A 和 B 两条链，还有两个蛋白分别为 L 链和 R 链。

隐藏 label 直接输入 Label 即可。



分别选择 DNA 和两个蛋白，并命名为 dna,prol,pror

```
select dna,chain a or chain b
```

```
select prol,chain l
```

```
select pror,chain r
```

隐藏所有显示模式

```
hide all
```

设置 dna 的显示模式为 sticks，设置 DNA 中的 P 原子为 spheres 和 cartoon 模式。

```
show sticks,dna
```

```
show cartoon,name p
```

```
show spheres,name p
```

设置两个蛋白为 cartoon 模式

```
show cartoon,prol or pror
```

改变两个蛋白的显示颜色

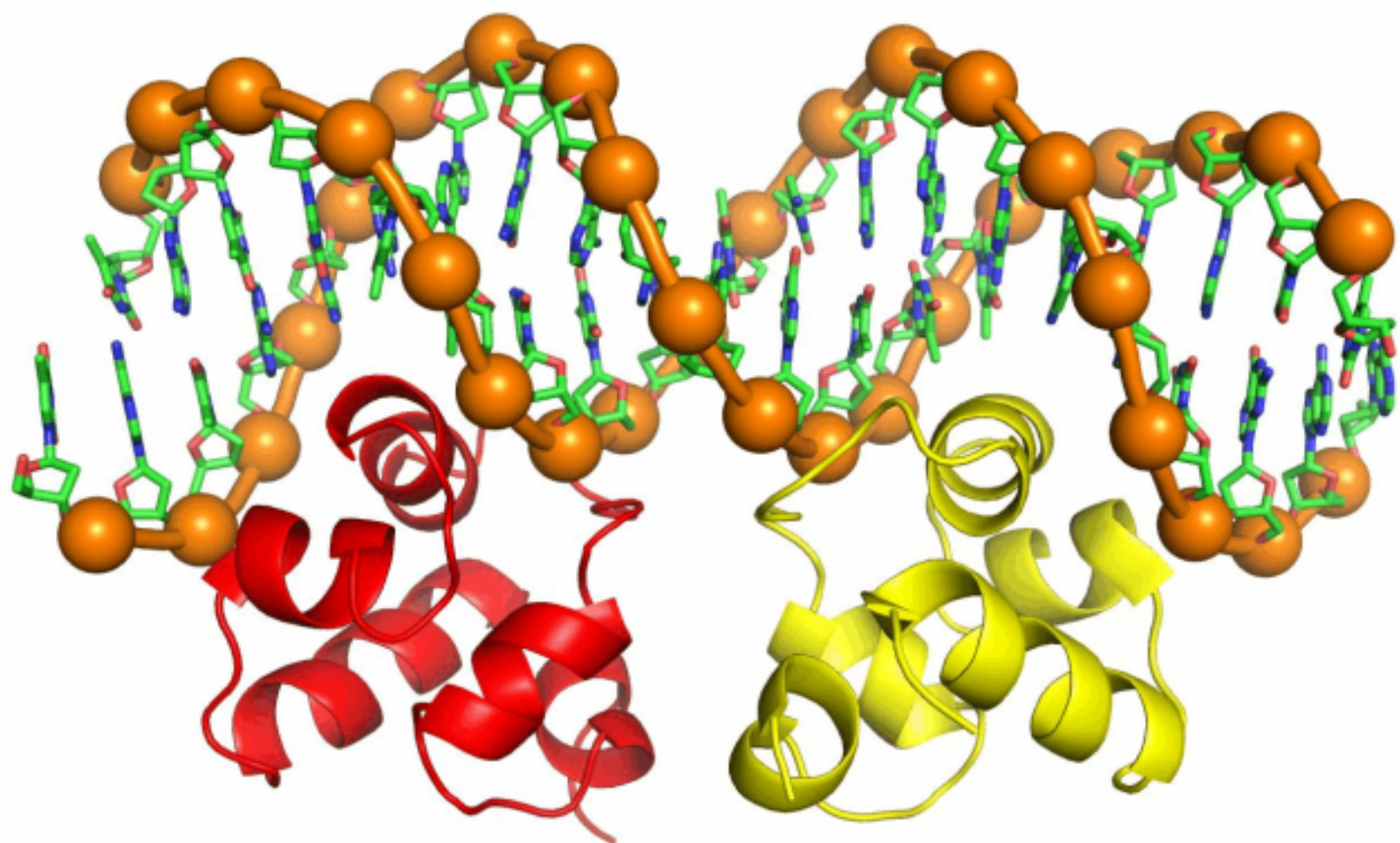
```
color red,prol
```

```
color yellow,pror
```

改变背景颜色为白色

```
bg_color white
```


此时是不是已经基本显示出下图的效果了？ 但好像还有点粗糙。另外，怎么把做好的图保存下来啊？难道要用截图工具截图吗？



结构打光并保存图片

```
ray
```

```
png 001
```

打光命令 `ray` 就像在照相馆照相时需要专门的灯光设备，这里 `ray` 打光后，图像显示出景深和立体效果，更加逼真和精美。

保存图片命令非常简单，`png` 后面加上你起的任意文件名即可，回车后图片自动保存在当前工作路径下，图片文件格式是 `png`，看看后缀就清楚了。注意当前工作路径在哪里，还记得 `pwd` 命令吧 O(_)O

高清大图怎么做呢？无非就是像素调高一些呗，还是 `ray` 命令，不过后面加上横轴和纵轴的像点数

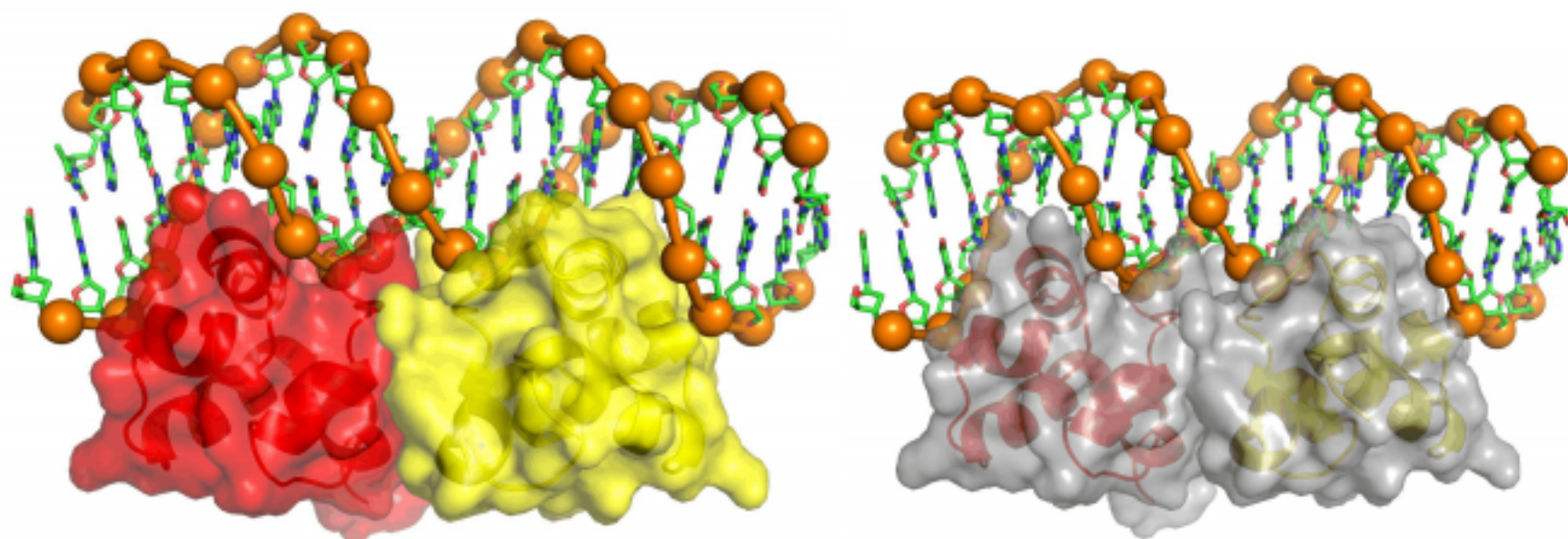
```
ray 2000,2000
```

```
png 002
```

看一看是不是高清多了，文件也大多了。如果 `ray` 后不加参数的话，那么就以 PyMOL 软件窗口的大小作图，你调的窗口大它就大，你调的小它就小。

好了，现在咱们已经完成第一张精美图片的制作了，很有成就感吧 (^ ^)

接下来，咱们来做第二张和第三张图片

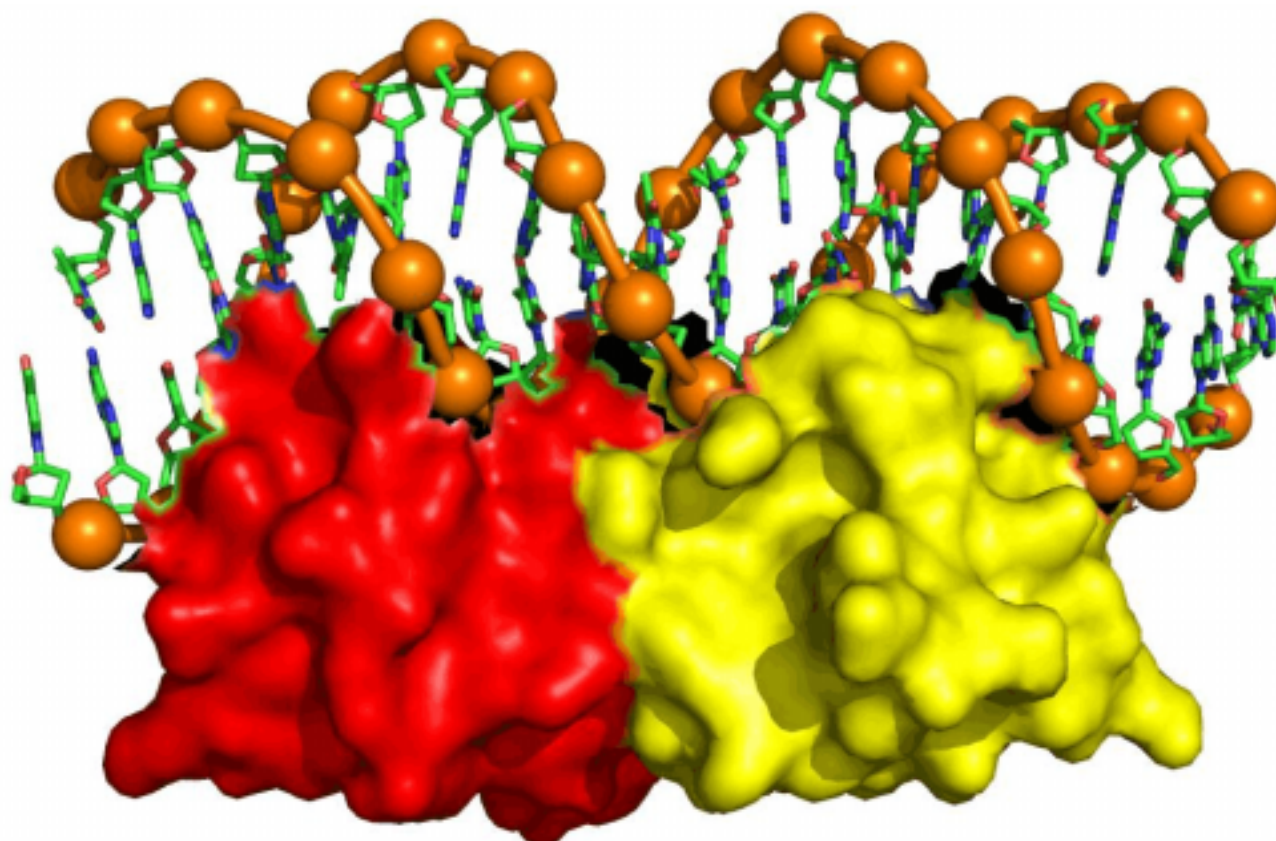


看到这两张图，可能你会觉得很简单，不就是把蛋白 show 出 surface 模式吗？可是如果真的这么做

```
show surface,prol
```

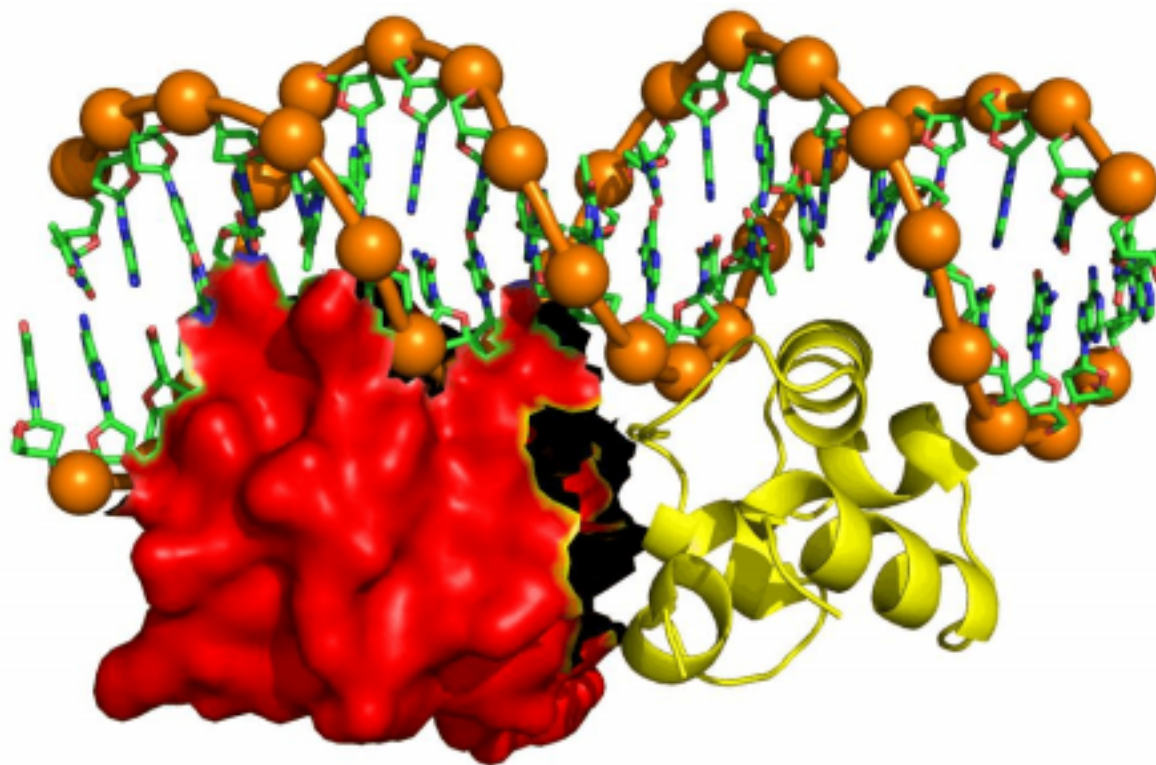
```
show surface,pror
```

会发现效果竟然是下面这张图



咦，这是什么情况？！怎么这个图里两个蛋白表面连到一起了？还有蛋白和 DNA 的交界面怎么会是裂开的啊？

还记得前面说过的 Object 和 selection 的区别吗？前面没理解的话，在这里会有深刻体会的。前面说过 Object 是实实在在的物体或东西，而 selection 不过是一个指代名称。选择链 l 为 prol，不过是用 prol 指代 3cro 中的链 l，但链 l 和链 r 还是在 3cro 中，而没有独立出来，两个蛋白链 l 与 r 和 DNA 链 a 与 b 共处在一个 object 中，它们是一体的，而不是独立的。既然大家都是一体的，显示表面时相互之间当然就是连接的，所以只显示蛋白的表面当然显示出裂开的蛋白 DNA 交界面，如果你只显示 prol 的话，还会出现裂开的蛋白交接面呐，见下图



那怎么显示出锁钥契合而不是相互连成一片的表面呢？既然大家共处一个 object 办不到，那就独立门户各成一体。建立 object 使用下面的命令

```
create pro1,prol
```

```
create pro2,pror
```

上面命令的意思就是创建一个名为 pro1 的 object，这个 object 的原子坐标取自 3cro 中的 prol。

建好 object 了，就分别显示蛋白的表面

```
show surface,pro1
```

```
show surface,pro2
```

这一次是不是不再连成一片了？

最后设置一下 surface 的透明度，把表面内的 cartoon 显示出来

```
set transparency,0.4
```

透明度取值从 0 到 1，0 是完全不透明，1 是完全透明。

到此再 ray 一下，保存即可得到第二张精美图。

第三张图只是把表面的颜色改了改，如果用前面所学的 color 改的话，连 cartoon 的颜色也会改掉，而只想改变 surface 的颜色，得使用下面的命令

```
set surface_color, gray70,pro1 or pro2
```

把 pro1 和 pro2 的表面颜色改为灰 70，gray70 配合白色背景非常好。也可以只改 pro1 或 pro2 的表面颜色。ray 一下，保存即可得到第三张精美图。

好啦，总算做完这三张图了 \(^o^)/

最后再补充一点，这么精美的构图打了这么半天命令才打出来，能不能把这

个效果用软件保存下来，以后直接打开这个保存文件就能恢复呢？

当然可以了，不知你有没有发现，PyMOL是没有撤销功能的，而只能通过不断的保存中间文件来代替撤销功能。怎么保存呢？菜单栏 `file` 里有个 `save session`，点击并起个文件名保存为 `pse` 格式即可。双击这个 `pse` 文件就能恢复。

3. PyMOL 学习

上面介绍的只是 PyMOL 的基本命令和操作，很多功能都没有涉及，比如距离、角度及二面角的测量、静电势分布计算、氢键显示、蛋白结构编辑、突变残基、动画制作以及 python 编程等等。这些功能可随用随学，不必一下子全学会。如果感兴趣，可继续学习《pymol 教程》。当然，别忘了 [PyMOLwiki](#) 网站！