

General Suggestions for Using MacroDox

MacroDox comes compiled for different size systems. The executables in the MacroDox directory are named according to the upper limit for the number of atoms it can treat. It is more efficient to use the version that fits your particular molecule's needs. The larger versions of MacroDox require so much memory that some functions are no longer functional (look for notes concerning this within the tutorials).

The different executables found in /disk06/macrodox3.0.0/program/ are

```
macrodox5K for 5000 or fewer atoms
macrodox12K for 12000 or fewer atoms
macrodox18K for 18000 or fewer atoms
macrodox21K for 21000 or fewer atoms
```

The tutorial is written using a small protein as an example and will function with the 5K version.

For novice Unix users, I recommend creating aliases in your .cshrc file in your personal user directory (e.g., mine is /usr/people/kathryn/). These aliases appear at the bottom of the file. Each time you login, this file will be noted so that you don't have to type long commands to reach whatever version of MacroDox you want.

Examples from my .cshrc file

```
alias mdx5k /disk06/macrodox3.0.0/program/macrodox5K
alias mdx12k /disk06/macrodox3.0.0/program/macrodox12K
alias mdx18k /disk06/macrodox3.0.0/program/macrodox18K
alias mdx21k /disk06/macrodox3.0.0/program/macrodox21K
```

If you wish to use these yourself, edit your .cshrc file using the jot editor and add these lines to the bottom of the file. Save the changes. Then logout and back in; this will activate the changes you made. Then when you want to access the version for 5000 atoms all you need to type is mdx5k. If you want the version for 21000 atoms all you need to type is mdx21k.

Table of Contents for Tutorials

Tutorial I	Importing Proteins into MacroDox and Assigning Charges	2
Tutorial II	Calculating the Electrostatic Field around a Protein	31
Tutorial III	Preparing a Second Protein for BD	44
Tutorial IV	Setting Up and Running BD Simulations	53
Tutorial V	Analyzing a BD Simulation	70
Tutorial VI	Using the BD Helix Algorithm	90

-----TUTORIAL SESSION I-----v3.0.0--

Importing Proteins into MacroDox and Assigning Charges

The following is an example tutorial session which reads in a Protein Data Bank file for the protein yeast iso-1-cytochrome c (lycc) and assigns the charges to every atom based on a Tanford-Kirkwood calculation at a specified pH, temperature and ionic strength.

Basically the types of operations exemplified here can be summarized as follows:

- 1) A Protein Data Bank file is imported using the NEW command. Default charge assignments are made from the 'macprep.chg' file of charges.
- 2) The file of atoms is edited, the waters are eliminated and the resultant file is saved as an 'mdx' formatted file using SAVE.
- 3) The status of charge assignments and the global properties of the molecule are looked at using the STATUS and PROPERTIES command, respectively.
- 4) A Tanford-Kirkwood calculation is set up and run using the TK menu of commands.
- 5) The ASSIGN command is used to implement these TK-recommended set of charges.
- 6) The ADJUST command is used to tweak the charge on the FE to be what it should be.
- 7) Finally, the PROPERTIES command is used for the properly assigned set of charges and a list of atoms near the resultant dipole vector is output to the file 'DIPOLE.LIS'.

1.1 Importing a Protein Data Bank File Using the NEW Command.

The program is initiated by typing the command

mdx5k

which brings up the main menu:

```

-----MacroDox Main Menu-----
|   File       Charges   Modify   Calculate   Dock       |
|-----|-----|-----|-----|-----|
|   new        assign    addh     properties  init       |
|   open       adjust    delh     surface     search     |
|   save       tk         mutate   field       rate       |
|   save as    status    xtrans   contour     review     |
|   swap                               edit       slice     |
|   help                               compare    |
|   settings                               renumber  |
|   quit                                           |
|-----|-----|-----|-----|-----|
-----atoms-----
Current file 1 = untitled           0
Current file 2 = untitled           0
Terminal type = IRIS workstation
-----v3.0.1

```

To import a brand new coordinate file for ycc in PDB format, use the NEW command. Commands can be all upper case, or all lower case, but not a mixture.

NEW lycc.pdb

The NEW command does the following:

-asks for a file name and then opens an ascii Protein Data Bank (PDB) formatted protein coordinate file called lycc.pdb

-a list of alternate atom names in a file called 'map.list' is consulted in order to rename certain atoms with other more acceptable names.

-it then assigns masses and charges to all identifiable amino acid residues from the default charge file 'macprep.chg'. These are charge assignments in which every atom is assigned a partial charge from the CHARMM16 charge set.

-alternative charge assignments can be made subsequently through the ASSIGN menu, which allows different Charge files to be attached in place of the default 'macprep.chg'.

-an alert will appear for any amino acid residue or hetero atoms not assignable through the currently attached charge file. A list of these is generated in the temporary file named 'NOASSIGN.LIS' and an edit window opens up with this file.

-titratable amino acids are identified for later calculation in Tanford-Kirkwood analysis. (TK)

-informs the user of the number of atoms read in, and the current position of the center of mass. If the center of mass is not at the origin, the user will be asked if a shift of the center of mass to the origin is desired.

-the prospective file generated, which is called the current MDX file is designated with the default name 'untitled' until SAVE or SAVE AS is invoked.

-MDX is a special formatted file which is identical to the PDB format through the atom coordinates, but also contains mass, charge assignment, surface accessibility, and other titration site information.

-a PDB file opened by NEW can be subject to modification through other commands and the resulting MDX file subsequently saved.

After the NEW lycc.pdb command, the program should return the following messages:

```
Number of atoms originally read in =          1014

HEM E    name changed to FE
HEM N A  name changed to NA
HEM N B  name changed to NB
HEM N C  name changed to NC
HEM N D  name changed to ND
Warning:          121 atom(s) not found in charge_table file:
Consult the output file NOASSIGN.LIS for a listing
of atoms with unassigned charge and mass.
Present position of center of mass =   4.507  19.824   7.983
```

Do you want to shift computed center of mass to the origin?

To the question "Do you want to shift computed center of mass to the origin?", answer **yes**. Following this the program advises the user where the center of mass is:

```
New position of center of mass =   0.000   0.000   0.000
```

and then returns to the MacroDox Main Menu.

Also, an edit window should open up containing the list of 121 atoms whose charges could not be assigned. This is the file named 'NOASSIGN.LIS'.

Here is a sample of that window containing the list of atoms that were not identified in the file 'macprep.chg', and thus have unassigned masses and charges. Here is a short sample of what that file should contain:

```
HETATM 895 O HOH 105
HETATM 896 O HOH 106
HETATM 897 O HOH 107
HETATM 898 O HOH 108
HETATM 899 O HOH 109
HETATM 900 O HOH 110
HETATM 901 O HOH 111
HETATM 902 O HOH 112
HETATM 903 O HOH 113
HETATM 904 O HOH 114
HETATM 905 O HOH 115
HETATM 906 O HOH 116
HETATM 907 S SO4 117
HETATM 908 O1 SO4 117
HETATM 909 O2 SO4 117
HETATM 910 O3 SO4 117
HETATM 911 O4 SO4 117
HETATM 912 O HOH 118
. . . . .
```

After inspecting this file, we find that these atoms are waters and a sulfate counterion. These are of no consequence to our subsequent treatment. After we **exit the 'NOASSIGN.LIS' windows**, we should go into the EDIT function and simply delete these from the newly created MDX file. But before we do that, it would not hurt to go ahead and save the assignments that have already been made. Execute the command:

SAVE

The program will ask you what type of file you want to create, another Protein Data Bank formatted file similar to what you imported with the NEW command, or whether you want the more information-expanded Macrodox-formatted file containing the charge assignments, etc. To this question, answer "mdx".

Do you want an MDX or PDB format file?

MDX

What name shall I give the MDX file?

lycc.mdx

Since we have done a SAVE, we could QUIT the program at this point. Let's do that and then restart it:

QUIT

1.2 The MDX File.

Here is what the MDX file we just created looks like, just showing the first 50 records, followed by annotation which describes the structure of this centrally important data base:

The first line is NATOM. The next lines contain the following variables in order:

speci	inumber	aname	resname	q	iresnum	xv(1)	xv(2)	xv(3)	charge	ans	distance	tchg	actg	pkint	bur	tac
	1014															
ATOM	1	N	THR	-5		0.246	-7.845	-16.955	-0.080	15.015	18.683	1.000	0.000	8.000	0.550	0.000
ATOM	2	CA	THR	-5		0.032	-8.374	-15.542	0.060	13.019	17.654	0.000	0.000	0.000	0.550	0.000
ATOM	3	C	THR	-5		-1.461	-8.758	-15.526	0.360	12.011	17.885	0.000	0.000	0.000	0.550	0.000
ATOM	4	O	THR	-5		-1.924	-9.461	-16.453	-0.363	15.999	19.076	0.000	0.000	0.000	0.550	0.000
ATOM	5	CB	THR	-5		0.523	-7.333	-14.540	0.140	13.019	16.293	0.000	0.000	0.000	0.550	0.000
ATOM	6	OG1	THR	-5		-0.154	-6.081	-14.622	-0.130	17.007	15.837	0.000	0.000	0.000	0.550	0.000
ATOM	7	CG2	THR	-5		2.011	-7.017	-14.666	0.013	15.035	16.382	0.000	0.000	0.000	0.550	0.000
ATOM	8	N	GLU	-4		-2.151	-8.185	-14.576	-0.095	15.015	16.854	0.000	0.000	0.000	0.550	0.000
ATOM	9	CA	GLU	-4		-3.622	-8.267	-14.395	0.050	13.019	16.990	0.000	0.000	0.000	0.550	0.000
ATOM	10	C	GLU	-4		-3.964	-6.928	-13.709	0.355	12.011	15.863	0.000	0.000	0.000	0.550	0.000
ATOM	11	O	GLU	-4		-4.921	-6.792	-12.914	-0.380	15.999	15.398	0.000	0.000	0.000	0.550	0.000
ATOM	12	CB	GLU	-4		-4.013	-9.424	-13.518	0.010	14.027	16.960	0.000	0.000	0.000	0.550	0.000
ATOM	13	CG	GLU	-4		-3.549	-10.817	-13.994	-0.155	14.027	18.040	0.000	0.000	0.000	0.550	0.000
ATOM	14	CD	GLU	-4		-4.083	-11.967	-13.190	0.365	12.011	18.272	0.000	0.000	0.000	0.550	0.000
ATOM	15	OE1	GLU	-4		-3.380	-12.946	-13.045	-0.575	15.999	18.687	-1.000	0.000	4.500	0.550	0.000
ATOM	16	OE2	GLU	-4		-5.294	-11.720	-12.735	-0.575	15.999	18.099	-1.000	0.000	4.500	0.550	0.000
ATOM	17	N	PHE	-3		-3.099	-5.978	-14.014	-0.080	15.015	15.547	0.000	0.000	0.000	0.550	0.000
ATOM	18	CA	PHE	-3		-2.971	-4.617	-13.533	0.075	13.019	14.604	0.000	0.000	0.000	0.550	0.000
ATOM	19	C	PHE	-3		-2.558	-3.745	-14.733	0.350	12.011	15.415	0.000	0.000	0.000	0.550	0.000
ATOM	20	O	PHE	-3		-1.522	-4.165	-15.339	-0.365	15.999	15.967	0.000	0.000	0.000	0.550	0.000
ATOM	21	CB	PHE	-3		-1.785	-4.515	-12.500	0.005	14.027	13.409	0.000	0.000	0.000	0.550	0.000
ATOM	22	CG	PHE	-3		-1.402	-3.092	-12.333	0.039	12.011	12.791	0.000	0.000	0.000	0.550	0.000
ATOM	23	CD1	PHE	-3		-2.255	-2.225	-11.627	-0.019	13.019	12.050	0.000	0.000	0.000	0.550	0.000
ATOM	24	CD2	PHE	-3		-0.283	-2.565	-12.939	-0.019	13.019	13.194	0.000	0.000	0.000	0.550	0.000
ATOM	25	CE1	PHE	-3		-1.965	-0.872	-11.502	0.010	13.019	11.701	0.000	0.000	0.000	0.550	0.000
ATOM	26	CE2	PHE	-3		0.015	-1.201	-12.885	0.010	13.019	12.941	0.000	0.000	0.000	0.550	0.000

speci	inumber	aname	resname	q	iresnum	xv(1)	xv(2)	xv(3)	charge	ans	distance	tchg	achg	pkint	bur	tac
ATOM	27	CZ	PHE	-3	-0.859	-0.348	-12.146	-0.006	13.019	12.181	0.000	0.000	0.000	0.550	0.000	
ATOM	28	N	LYS	-2	-3.212	-2.621	-14.949	-0.090	15.015	15.513	0.000	0.000	0.000	0.550	0.000	
ATOM	29	CA	LYS	-2	-2.855	-1.720	-16.050	0.080	13.019	16.392	0.000	0.000	0.000	0.550	0.000	
ATOM	30	C	LYS	-2	-2.857	-0.304	-15.476	0.348	12.011	15.740	0.000	0.000	0.000	0.550	0.000	
ATOM	31	O	LYS	-2	-3.628	-0.030	-14.546	-0.360	15.999	14.991	0.000	0.000	0.000	0.550	0.000	
ATOM	32	CB	LYS	-2	-3.780	-1.716	-17.233	0.040	14.027	17.726	0.000	0.000	0.000	0.550	0.000	
ATOM	33	CG	LYS	-2	-3.528	-2.869	-18.226	0.054	14.027	18.784	0.000	0.000	0.000	0.550	0.000	
ATOM	34	CD	LYS	-2	-4.289	-2.664	-19.561	0.063	14.027	20.202	0.000	0.000	0.000	0.550	0.000	
ATOM	35	CE	LYS	-2	-5.794	-2.692	-19.400	0.205	14.027	20.424	0.000	0.000	0.000	0.550	0.000	
ATOM	36	NZ	LYS	-2	-6.288	-4.077	-19.141	0.660	17.031	20.555	1.000	0.000	10.400	0.550	0.000	
ATOM	37	N	ALA	-1	-2.033	0.524	-16.103	-0.080	15.015	16.239	0.000	0.000	0.000	0.550	0.000	
ATOM	38	CA	ALA	-1	-1.838	1.912	-15.668	0.075	13.019	15.890	0.000	0.000	0.000	0.550	0.000	
ATOM	39	C	ALA	-1	-3.140	2.715	-15.803	0.350	12.011	16.339	0.000	0.000	0.000	0.550	0.000	
ATOM	40	O	ALA	-1	-3.879	2.504	-16.789	-0.365	15.999	17.412	0.000	0.000	0.000	0.550	0.000	
ATOM	41	CB	ALA	-1	-0.736	2.658	-16.382	0.020	15.035	16.612	0.000	0.000	0.000	0.550	0.000	
ATOM	42	N	GLY	1	-3.227	3.701	-14.908	-0.080	15.015	15.695	0.000	0.000	0.000	0.550	0.000	
ATOM	43	CA	GLY	1	-4.347	4.587	-14.839	0.090	14.027	16.128	0.000	0.000	0.000	0.550	0.000	
ATOM	44	C	GLY	1	-3.772	5.970	-14.772	0.345	12.011	16.373	0.000	0.000	0.000	0.550	0.000	
ATOM	45	O	GLY	1	-2.719	6.358	-15.355	-0.355	15.999	16.840	0.000	0.000	0.000	0.550	0.000	
ATOM	46	N	SER	2	-4.471	6.796	-14.051	-0.080	15.015	16.235	0.000	0.000	0.000	0.550	0.000	
ATOM	47	CA	SER	2	-4.198	8.198	-13.867	0.060	13.019	16.647	0.000	0.000	0.000	0.550	0.000	
ATOM	48	C	SER	2	-3.434	8.337	-12.533	0.360	12.011	15.439	0.000	0.000	0.000	0.550	0.000	
ATOM	49	O	SER	2	-4.015	8.217	-11.441	-0.363	15.999	14.647	0.000	0.000	0.000	0.550	0.000	
ATOM	50	CB	SER	2	-5.499	9.013	-13.815	0.130	14.027	17.387	0.000	0.000	0.000	0.550	0.000	

The FORTRAN format of these records are as follows:

```
a6,i5,2x,a4,a4,a1,i4,4x,11f8.3
```

Here are the definitions of these variables:

NATOM = total number of atoms in the molecule.

SPECI(I) = species type of atom I. (ATOM, HETATM, TER, etc.)

INUMBER(I) = atom sequence number of atom I. (a number from 1 to NATOM)

ANAME(I) = atom name of atom I. (CA, C, N, O, CB, etc.)

RESNAME(I) = residue name of atom I. (GLY, HIS, PRO, etc.)

Q(I) = segment designator of atom I. (for multi-subunit proteins)

IRESNUM(I) = residue number of atom I.

XV(K,I) = coordinates of atom I. (K= 1,2,3 = x,y,z positions of the molecule)

CHARGE(I)= currently assigned electrostatic charge on atom I. This is initially assigned by the CHARGE_TABLE subroutine using the partial atomic charge set found in the parameter file 'macprep.chg'. This set may be modified to contain formal charges using the ASSIGN subroutine, or may inherit charges recommended following a Tanford-Kirkwood calc, or may be manually changed by the user at will in edit sessions.

AMS(I) = atomic mass of atom I in amu. These are assigned upon reading in a new coordinate file with the NEW command. This is done by the ASSIGN subroutine. Mass values are determined by the parameter file 'macprep.chg'.

DISTANCE(I) = distance of atom I from center of mass of the molecule. These are computed by the subroutine DIST when reading in a new coordinate set with the NEW command.

TCHG(I) = maximum formal charge on atom I for Tanford-Kirkwood calculation.

ACHG(I) = recommended mean charge on atom I after Tan-Kirk calculation.

PKINT(I) = intrinsic pKa of a titratable site. (0 for other atoms)

BUR(I) = burial ratio of atom I. (only those with TCHG not equal to 0)

TAC(I) = surface accessibility of atom I after Richards' or Quik surface cal.

Here is the information at the end of the atom records of the MDX file followed by their definitions:

```

504.654694          dipole magnitude
-452.647888  220.622803  -33.343990      dipole vector
  1                if heme is present
848                monitor atom seq# (Fe by default)
849                | set of 4 'reactive' atoms
850                | heme peripheral atom seq #s
851                | here by default
852                |
103                res # bearing the oxy terminus
-0.628336  -0.776256  -0.051193      heme plane normal (if HEM present)
-0.000009  -0.000027  -0.000016      center of mass
-0.475480   0.065167  -0.531111      center of + charge
  0.742650  -0.540278  -0.476288      center of - charge
80.613998          total + charge
75.314056          total - charge
23.701630          farthest dist from c. of m.
13.274193          radius of gyration
17.500000          est. effective spherical radius
  0.012457          translational diffusion coeff
  0.000031          rotational diffusion coeff

```

Here are these variable names in the same order:

DPL = dipole moment magnitude corresponding to charges in array CHARGE.

DIP(1..3) = dipole moment vector corresponding to charges in array CHARGE.

IHEME = flag to say whether the protein is a heme protein. (1=Yes)

NFE = atom sequence number of FE atom in a heme-containing protein, or the "monitor atom" in general. (in BD runs)

NCHA, NCHB, NCHC, NCHD = atom sequence numbers of heme peripheral atoms in a heme-containing protein, or the "reactive atom set" in general. (in BD runs)
 IRESOX=residue number containing the OXY terminus.
 TI(1..3) = normal vector of heme plane, if applicable.
 CM(1..3) = center of mass of molecule.
 RP(1..3) = center of positive charge.
 RN(1..3) = center of negative charge.
 TP = total positive charge.
 TN = total negative charge.
 GREATEST = distance from center of mass to atom furthest out.
 RGYR = radius of gyration of molecule.
 RAD = estimated spherical radius of the molecule.
 DTRANS = Stokes-Einstein estimated translational diffusion coefficient.
 DROT = Stokes-Einstein estimated rotational diffusion coefficient.

1.3 The EDIT, PROPERTIES, and STATUS Commands.

 Now let us restart the program with

mdx5k

and open the newly-created Macrodox-formatted protein file with the command:

OPEN lycc.mdx

Notice that the main menu lists the Current file name as follows:

```

-----atoms-----
Current file 1 = lycc.mdx                1014
Current file 2 = untitled                0
Terminal type = IRIS workstation
-----v3.0.0

```

Now, use the EDIT function

EDIT

to initiate a 'jot' editing session on the MDX file. Here, we go into this MDX file and delete *the unnecessary solvent atoms and the three methyl groups of TML72*. We will simply redefine this residue as

Lysine for convenience here, although we do have the TML residue defined in the 'macprep.chg' file. **Delete the HETATM atoms beginning with number 891 and going through 1014.** Note: 891-893 are tge TML sidechain carbons. 894-1014 are O atoms of water. Do not delete the heme, which are the preceding HETATM. Do not delete the special numerical info at the end of the MDX file. Also notice that the porphyrin residue HEM is given a residue #1. We choose to **renumber this as residue 104 manually using the editor.**

A partial listing of the edited MDX file should look like this:

Notice that the number of atoms at the top of the file needs to be changed from 1014 to the new number of atoms remaining after the edit. MacroDox will automatically recount the atoms and replace this number with the correct number, 890.

```

      890
ATOM   1  N  THR   -5    0.246  -7.845 -16.955  -0.080  15.015  18.683  1.000  0.000  8.000  0.550  0.000
ATOM   2  CA THR   -5    0.032  -8.374 -15.542   0.060  13.019  17.654  0.000  0.000  0.000  0.550  0.000
ATOM   3  C  THR   -5   -1.461  -8.758 -15.526   0.360  12.011  17.885  0.000  0.000  0.000  0.550  0.000
ATOM   4  O  THR   -5   -1.924  -9.461 -16.453  -0.363  15.999  19.076  0.000  0.000  0.000  0.550  0.000
ATOM   5  CB THR   -5    0.523  -7.333 -14.540   0.140  13.019  16.293  0.000  0.000  0.000  0.550  0.000
ATOM   6  OG1 THR  -5   -0.154  -6.081 -14.622  -0.130  17.007  15.837  0.000  0.000  0.000  0.550  0.000
ATOM   7  CG2 THR  -5    2.011  -7.017 -14.666   0.013  15.035  16.382  0.000  0.000  0.000  0.550  0.000
ATOM   8  N  GLU   -4   -2.151  -8.185 -14.576  -0.095  15.015  16.854  0.000  0.000  0.000  0.550  0.000
ATOM   9  CA GLU   -4   -3.622  -8.267 -14.395   0.050  13.019  16.990  0.000  0.000  0.000  0.550  0.000
ATOM  10  C  GLU   -4   -3.964  -6.928 -13.709   0.355  12.011  15.863  0.000  0.000  0.000  0.550  0.000
ATOM  11  O  GLU   -4   -4.921  -6.792 -12.914  -0.380  15.999  15.398  0.000  0.000  0.000  0.550  0.000
ATOM  12  CB GLU   -4   -4.013  -9.424 -13.518   0.010  14.027  16.960  0.000  0.000  0.000  0.550  0.000
ATOM  13  CG GLU   -4   -3.549 -10.817 -13.994  -0.155  14.027  18.040  0.000  0.000  0.000  0.550  0.000
ATOM  14  CD GLU   -4   -4.083 -11.967 -13.190   0.365  12.011  18.272  0.000  0.000  0.000  0.550  0.000
ATOM  15  OE1 GLU  -4   -3.380 -12.946 -13.045  -0.575  15.999  18.687  -1.000  0.000  4.500  0.550  0.000
ATOM  16  OE2 GLU  -4   -5.294 -11.720 -12.735  -0.575  15.999  18.099  -1.000  0.000  4.500  0.550  0.000
.
.
.
ATOM  847 OXT GLU  103  11.257   7.595  -5.975  -1.000  15.999  14.836  -1.000  0.000  3.500  0.550  0.000
HETATM 848 FE  HEM  104  -1.703   1.153   5.220   0.240  55.847   5.611   0.000  0.000  13.000  0.550  0.000
HETATM 849 CHA HEM  104   0.599  -1.122   6.254   0.040  13.019   6.382   0.000  0.000  0.000  0.550  0.000
HETATM 850 CHB HEM  104  -0.537   0.827   2.047   0.040  13.019   2.272   0.000  0.000  0.000  0.550  0.000
HETATM 851 CHC HEM  104  -4.298   2.985   4.158   0.040  13.019   6.684   0.000  0.000  0.000  0.550  0.000
HETATM 852 CHD HEM  104  -2.402   1.910   8.374   0.040  13.019   8.919   0.000  0.000  0.000  0.550  0.000
HETATM 853 NA  HEM  104  -0.276   0.082   4.385  -0.180  14.007   4.395   0.000  0.000  0.000  0.550  0.000
HETATM 854 C1A HEM  104   0.594  -0.798   4.905   0.030  12.011   5.005   0.000  0.000  0.000  0.550  0.000
HETATM 855 C2A HEM  104   1.287  -1.592   3.862  -0.020  12.011   4.372   0.000  0.000  0.000  0.550  0.000
HETATM 856 C3A HEM  104   0.966  -1.004   2.709   0.020  12.011   3.047   0.000  0.000  0.000  0.550  0.000
HETATM 857 C4A HEM  104   0.009  -0.013   2.934   0.020  12.011   2.934   0.000  0.000  0.000  0.550  0.000
HETATM 858 CMA HEM  104   1.468  -1.450   1.283  -0.040  15.035   2.430   0.000  0.000  0.000  0.550  0.000
HETATM 859 CAA HEM  104   2.303  -2.642   4.210   0.040  14.027   5.479   0.000  0.000  0.000  0.550  0.000
HETATM 860 CBA HEM  104   3.708  -1.970   4.343  -0.100  14.027   6.041   0.000  0.000  0.000  0.550  0.000
HETATM 861 CGA HEM  104   4.924  -2.876   4.463   0.300  12.011   7.242   0.000  0.000  0.000  0.550  0.000
HETATM 862 O1A HEM  104   5.990  -2.263   4.591  -0.500  15.999   7.880  -1.000  0.000  4.000  0.550  0.000

```

HETATM	863	O2A	HEM	104	4.795	-4.119	4.444	-0.500	15.999	7.728	-1.000	0.000	4.000	0.550	0.000
HETATM	864	NB	HEM	104	-2.315	1.805	3.428	-0.180	14.007	4.513	0.000	0.000	0.000	0.550	0.000
HETATM	865	C1B	HEM	104	-1.565	1.736	2.260	0.030	12.011	3.251	0.000	0.000	0.000	0.550	0.000
HETATM	866	C2B	HEM	104	-2.354	2.431	1.247	0.020	12.011	3.606	0.000	0.000	0.000	0.550	0.000
HETATM	867	C3B	HEM	104	-3.458	2.996	1.796	-0.050	12.011	4.915	0.000	0.000	0.000	0.550	0.000
HETATM	868	C4B	HEM	104	-3.355	2.618	3.180	0.020	12.011	5.312	0.000	0.000	0.000	0.550	0.000
HETATM	869	CMB	HEM	104	-1.829	2.594	-0.226	-0.040	15.035	3.182	0.000	0.000	0.000	0.550	0.000
HETATM	870	CAB	HEM	104	-4.532	3.877	1.214	0.030	13.019	6.086	0.000	0.000	0.000	0.550	0.000
HETATM	871	CBB	HEM	104	-5.679	3.109	0.537	-0.100	14.027	6.496	0.000	0.000	0.000	0.550	0.000
HETATM	872	NC	HEM	104	-3.085	2.242	6.147	-0.180	14.007	7.234	0.000	0.000	0.000	0.550	0.000
HETATM	873	C1C	HEM	104	-4.116	2.949	5.607	0.030	12.011	7.555	0.000	0.000	0.000	0.550	0.000
HETATM	874	C2C	HEM	104	-4.911	3.716	6.483	0.020	12.011	8.942	0.000	0.000	0.000	0.550	0.000
HETATM	875	C3C	HEM	104	-4.404	3.375	7.686	-0.050	12.011	9.479	0.000	0.000	0.000	0.550	0.000
HETATM	876	C4C	HEM	104	-3.224	2.537	7.483	0.020	12.011	8.534	0.000	0.000	0.000	0.550	0.000
HETATM	877	CMC	HEM	104	-6.087	4.593	6.180	-0.040	15.035	9.815	0.000	0.000	0.000	0.550	0.000
HETATM	878	CAC	HEM	104	-4.650	3.973	9.106	0.030	13.019	10.969	0.000	0.000	0.000	0.550	0.000
HETATM	879	CBC	HEM	104	-6.150	3.972	9.478	-0.100	14.027	11.976	0.000	0.000	0.000	0.550	0.000
HETATM	880	ND	HEM	104	-1.029	0.527	7.000	-0.180	14.007	7.095	0.000	0.000	0.000	0.550	0.000
HETATM	881	C1D	HEM	104	-1.450	0.968	8.249	0.030	12.011	8.431	0.000	0.000	0.000	0.550	0.000
HETATM	882	C2D	HEM	104	-0.761	0.288	9.245	0.020	12.011	9.281	0.000	0.000	0.000	0.550	0.000
HETATM	883	C3D	HEM	104	0.063	-0.593	8.655	-0.020	12.011	8.676	0.000	0.000	0.000	0.550	0.000
HETATM	884	C4D	HEM	104	-0.162	-0.497	7.215	0.020	12.011	7.234	0.000	0.000	0.000	0.550	0.000
HETATM	885	CMD	HEM	104	-0.981	0.451	10.818	-0.040	15.035	10.872	0.000	0.000	0.000	0.550	0.000
HETATM	886	CAD	HEM	104	0.917	-1.655	9.392	0.040	14.027	9.581	0.000	0.000	0.000	0.550	0.000
HETATM	887	CBD	HEM	104	0.312	-3.038	9.331	-0.100	14.027	9.818	0.000	0.000	0.000	0.550	0.000
HETATM	888	CGD	HEM	104	0.814	-4.084	10.309	0.300	12.011	11.119	0.000	0.000	0.000	0.550	0.000
HETATM	889	O1D	HEM	104	1.966	-3.911	10.791	-0.500	15.999	11.646	-1.000	0.000	4.000	0.550	0.000
HETATM	890	O2D	HEM	104	0.089	-5.100	10.401	-0.500	15.999	11.585	-1.000	0.000	4.000	0.550	0.000
.															
.															
.															

So now, do a **save/exit** from the EDIT session (do not use SAVE/AS to try to rename the file.)

Press RETURN key after file is saved

Now we are back at the main menu, and are ready to perform other functions. First as a check, let's look at the current properties of the molecule:

properties

Since we have deleted some atoms, the center of mass position has been shifted, and there is a facility in PROPERTIES to shift atoms to bring the new center of mass to the origin. The program will say:

Present position of center of mass = 0.031 0.018 -0.042

Do you want to shift computed center of mass to the origin?

Answer **'yes'** to this question. You will see the following appear on the screen:

New position of center of mass = 0.000 0.000 0.000

Now you will see the following table of global properties of this molecule appear in a 'jot' edit window for your inspection:

```

-----
                properties
-----
number of atoms=          890
Total mass= 12671.35 amu
Center of mass =    0.000  0.000  0.000
Center of +/- charge= -0.476  0.066 -0.535    0.711 -0.558 -0.434
Total +/- charge=  80.314  75.314 Net Charge=    5.000e
Dipole magnitude=   498.32 Debye
Dipole vec =-440.709 227.294 -49.322
Greatest distance from center of mass =   22.584 A
Full surface calculation not performed yet
Radius of gyration=   12.72
Atom # density in concentric shells of thickness=  3.0
  shell #  density
      1     35.4    2     41.7    3     54.0    4     62.6
      5     40.3    6     15.5    7     2.3    8     0.2
Estimated effective spherical radius =   16.8 A
Est Stokes-Einstein trans and rot diff coeff =
0.130E-01 A**2/ps, 0.345E-04 1/ps
Heme group found
Heme plane normal vector:   -0.628 -0.776 -0.051
Angle between dipole and heme normal =    78.07
Angle between dipole and Fe vector =    74.30

```

You may rename and/or **exit** this window at your discretion. It will be on your hard disk as the file PROPERTIES.OUT.

Back in the text window, the question appears:

Generate listing of atoms along dipole vector axes?

Answer **'no'** to this question, since we have further charge refinement to perform before we will have an accurate charge set.

Notice that the net charge of the molecule is currently 5.00e and the dipole moment is 498.32 Debye. The radius of gyration is 12.72 Angstroms. All distances are in Angstroms, by the way.

It is also instructive to see what the current charge assignments are by residue. This can be accomplished with the status command:

STATUS

The following Charge_Status_List window appears:

Current charge assignments							
Residue		Curr Assignments		TK-recommended			
-5	THR	0.000	0.00	50	ASP	-1.000	0.00
-4	GLU	-1.000	0.00	54	LYS	1.000	0.00
-2	LYS	1.000	0.00	55	LYS	1.000	0.00
4	LYS	1.000	0.00	60	ASP	-1.000	0.00
5	LYS	1.000	0.00	61	GLU	-1.000	0.00
11	LYS	1.000	0.00	66	GLU	-1.000	0.00
13	ARG	1.000	0.00	67	TYR	0.000	0.00
14	CYS	0.000	0.00	72	LYS	1.000	0.00
17	CYS	0.000	0.00	73	LYS	1.000	0.00
18	HIS	0.000	0.00	74	TYR	0.000	0.00
21	GLU	-1.000	0.00	79	LYS	1.000	0.00
22	LYS	1.000	0.00	86	LYS	1.000	0.00
26	HIS	0.000	0.00	87	LYS	1.000	0.00
27	LYS	1.000	0.00	88	GLU	-1.000	0.00
33	HIS	0.000	0.00	89	LYS	1.000	0.00
38	ARG	1.000	0.00	90	ASP	-1.000	0.00
39	HIS	0.000	0.00	91	ARG	1.000	0.00
44	GLU	-1.000	0.00	93	ASP	-1.000	0.00
46	TYR	0.000	0.00	97	TYR	0.000	0.00
48	TYR	0.000	0.00	99	LYS	1.000	0.00
				100	LYS	1.000	0.00
				102	CYS	0.000	0.00
				103	GLU	-2.000	0.00
				104	HEM	-2.000	0.00
-----				-----			
Tot charge = 5.000				TK tot charge = 0.000			
-----				-----			

The right-hand column of charges are 0.00 because no Tanford-Kirkwood (TK) calculation has yet been performed and saved on this molecule. This column is the TK-recommended charges. The third column is the currently assigned net charge per residue.

Notice that

- (1) the N-terminal residue is uncharged. This may be inappropriate at the desired pH.
- (2) all acidic residues ASP and GLU bear a -1.00 charge. Again this might be inappropriate if you were working at low pH.
- (3) all basic residues LYS and ARG are +1.00 charge. This might be inappropriate at high pH.

- (4) The HIS residues are assumed to be unprotonated and have a charge of 0.00. This is certainly a zero-order approximation at neutral pH, since the intrinsic pK of HIS is 6.0
- (5) other titratable residues such as CYS and TYR are 0.00
- (6) Hetero atoms such as the FE atom may need their charges adjusted depending on their desired oxidation state. The HEM residue in our 'macprep.chg' file assumes a Fe(II) oxidation state, giving a formal charge on the Fe atom of 0.0. The net charge of HEM is thus -2.0 due to the two propionate groups. If you desire Fe(III) some adjustments will be necessary.

Exit the Charge_Status_List window, renaming it if you prefer, and type a carriage return in the text port to proceed.

Hit **RETURN** to continue

1.4 Doing a Tanford-Kirkwood Titration Calculation Using TK.

A Tanford-Kirkwood calculation should be performed to decide how to assign charges on a more refined level at the pH you desire. That is the next step in the operation.

The TK command:

-performs a Tanford-Kirkwood analysis on current MDX file.

This calculation estimates the pH and ionic strength dependent pKa's of ionization of individual amino acid residues and predicts the protonation status of titratable sites at various pH and ionic strength. It invokes a menu for stepwise TK analysis, including the following functions which are typically executed in the same sequence:

```

Display or Adjust Titratable Sites
Surface Accessibility Calculations
Review Close Ionic Contacts
Perform Tanford-Kirkwood Calculation
Return to Main Menu

```

Now let's invoke the TK command:

TK

This generates the following menu:

```

-----TK Menu-----
| SITES   - Adjust or Display Titratable Sites |

```

```

| BURIAL - Surface Accessibility Calculation
| INSPECT - Review Close Ionic Contacts
| TITRATE - Tanford-Kirkwood Calculation
| RETURN - Exit to Main Menu
| HELP
|-----

```

First we invoke the SITES command:

SITES

"Adjust or Display Titratable Sites" produces a table called sites.list showing the current list of titratable sites identified by the initial screening procedure invoked by NEW. At this juncture the user may choose to modify the sites in the 'jot' edit window produced by this SITES command:

Current list of titratable sites

```

-----
atom# atom res res#  chg  <pKint>  <burial>
-----
  1  N  THR   -5   1.0  8.00  0.55
 15  OE1 GLU   -4  -1.0  4.50  0.55
 16  OE2 GLU   -4  -1.0  4.50  0.55
 36  NZ  LYS   -2   1.0 10.40  0.55
 65  NZ  LYS    4   1.0 10.40  0.55
 74  NZ  LYS    5   1.0 10.40  0.55
118  NZ  LYS   11   1.0 10.40  0.55
135  NH1 ARG   13   1.0 12.00  0.55
136  NH2 ARG   13   1.0 12.00  0.55
142  SG  CYS   14  -1.0  9.00  0.55
165  SG  CYS   17  -1.0  9.00  0.55
172  ND1 HIS   18   1.0  6.30  0.55
175  NE2 HIS   18   1.0  6.30  0.55
197  OE1 GLU   21  -1.0  4.50  0.55
198  OE2 GLU   21  -1.0  4.50  0.55
207  NZ  LYS   22   1.0 10.40  0.55
229  ND1 HIS   26   1.0  6.30  0.55
232  NE2 HIS   26   1.0  6.30  0.55
241  NZ  LYS   27   1.0 10.40  0.55
282  ND1 HIS   33   1.0  6.30  0.55
285  NE2 HIS   33   1.0  6.30  0.55
322  NH1 ARG   38   1.0 12.00  0.55
323  NH2 ARG   38   1.0 12.00  0.55
330  ND1 HIS   39   1.0  6.30  0.55
333  NE2 HIS   39   1.0  6.30  0.55
365  OE1 GLU   44  -1.0  4.50  0.55
366  OE2 GLU   44  -1.0  4.50  0.55
382  OH  TYR   46  -1.0 10.00  0.55
400  OH  TYR   48  -1.0 10.00  0.55
414  OD1 ASP   50  -1.0  4.00  0.55
415  OD2 ASP   50  -1.0  4.00  0.55
445  NZ  LYS   54   1.0 10.40  0.55
454  NZ  LYS   55   1.0 10.40  0.55
498  OD1 ASP   60  -1.0  4.00  0.55
499  OD2 ASP   60  -1.0  4.00  0.55
507  OE1 GLU   61  -1.0  4.50  0.55
508  OE2 GLU   61  -1.0  4.50  0.55
546  OE1 GLU   66  -1.0  4.50  0.55
547  OE2 GLU   66  -1.0  4.50  0.55
559  OH  TYR   67  -1.0 10.00  0.55
598  NZ  LYS   72   1.0 10.40  0.55
607  NZ  LYS   73   1.0 10.40  0.55
619  OH  TYR   74  -1.0 10.00  0.55
654  NZ  LYS   79   1.0 10.40  0.55
703  NZ  LYS   86   1.0 10.40  0.55
712  NZ  LYS   87   1.0 10.40  0.55
720  OE1 GLU   88  -1.0  4.50  0.55
721  OE2 GLU   88  -1.0  4.50  0.55

```

730	NZ	LYS	89	1.0	10.40	0.55	837	SG	CYS	102	-1.0	9.00	0.55
737	OD1	ASP	90	-1.0	4.00	0.55	845	OE1	GLU	103	-1.0	4.50	0.55
738	OD2	ASP	90	-1.0	4.00	0.55	846	OE2	GLU	103	-1.0	4.50	0.55
748	NH1	ARG	91	1.0	12.00	0.55	847	OXT	GLU	103	-1.0	3.50	0.55
749	NH2	ARG	91	1.0	12.00	0.55	→848	FE	HEM	104	0.0	13.00	0.55
764	OD1	ASP	93	-1.0	4.00	0.55	862	O1A	HEM	104	-1.0	4.00	0.55
765	OD2	ASP	93	-1.0	4.00	0.55	863	O2A	HEM	104	-1.0	4.00	0.55
800	OH	TYR	97	-1.0	10.00	0.55	889	O1D	HEM	104	-1.0	4.00	0.55
817	NZ	LYS	99	1.0	10.40	0.55	890	O2D	HEM	104	-1.0	4.00	0.55
826	NZ	LYS	100	1.0	10.40	0.55							

Here it would be possible to use the editor to change various things pertaining to the subsequent titration calculation that will be performed. (Only the chg and <pKint> columns should be changed here.)

We desire this cytochrome to be in the Fe(III) oxidation state. We can accomplish that here by changing the chg value of FE atom from 0.0 to 1.0. Let's do that right now in this edit window.

Now we have atom # 848: FE HEM 104 1.0 13.00 0.55

Notice that the acidic residues such as GLU and ASP have two possible sites of -1 charge in this list. This will be remedied later by the BURIAL command under the TK menu, which will let the most exposed atom bear the full -1.0 charge.

Also notice that the burial ratio is 0.55 for all titratable sites. This will change after the BURIAL command has been executed.

Now save and exit this edit window. (don't change the file name)
Press **Return** to continue.

The next step in a Tanford-Kirkwood calculation is the Surface accessibility calculation.

This is invoked within the TK menu by the command:

BURIAL

'Surface Accessibility Calculations' does a Richards' solvent accessible surface analysis of every titratable groups formally charged atom only. To compute the surface accessibility of every atom, which is a much lengthier procedure, see the command SURFACE in the CALCULATE menu. Charges are placed on the most accesible formally-charged atom in ASP, GLU, ARG, HIS. A file 'SELECT.INF' is created containing info showing possible salt bridges. This may be examined by the next menu item INSPECT - Reviewing Close Contacts.

The following notice will appear on the screen:

```

-----
| Surface accessibility has been calculated for every titratable
| formally-charged atom only. Charges have been placed on the most
| accesible formally-charged atom in residues ASP, GLU, ARG, HIS.
| A file SELECT.INF has been created containing info showing possible
| salt bridges. This may be examined by the INSPECT command.
| Info in SELECT.INF may suggest to the user that switches in where
| charges have been placed be made in the MDX file.
| Also, disulphides involved in S-S crosslinks and metal ion ligands
| may need to be zeroed as well. These changes should be implemented
| in the MDX file prior to titrating by reissuing the SITES command.
| <RETURN> TO CONTINUE
-----

```

Now we are ready to review the close contacts determined by the surface accessibility calculation. Type the command:

INSPECT

An edit window entitled INSPECT.INF should appear on the screen which looks like this:

```

list of close ionic contacts found
res#   res name  charge      res# res  name  charge dist
  18   HIS  NE2   1.00 <-->  104  HEM  FE    1.00 1.99
->33   HIS  NE2   1.00 <-->  103  GLU  OXT  -1.00 3.28
  104   HEM  O1A   0.00 <-->   48  TYR  OH   -1.00 2.81

```

Notice here that HIS18 has been identified as ligated to the FE atom in the HEM residue, at a distance of 1.99. This implies that the HIS18 residue is ineligible from being in the protonated state, and we need to go back to the SITES command and zero out the possible charge assignable to HIS 18.

You may or may not wish to exit the window entitled SELECT.INF at this point. I would go ahead and **exit** it. It is already on disk.

Now in the TK Menu execute the SITES command again:

SITES

We see the edit window open up again with the list of titratable sites. Now it looks different because a BURIAL calculation has been performed and these no longer have their default values of 0.55. Also the formal charge has been placed on the least buried atom of the charged pairs on the sidechains of ASP, GLU and ARG.

```

Current list of titratable sites
-----
atom# atom res res#  chg  <pKint>  <burial>
-----
  1  N  THR   -5  1.0  8.00  0.33
 15 OE1 GLU   -4 -1.0  4.50  0.02
 16 OE2 GLU   -4  0.0  4.50  0.21
 36 NZ  LYS   -2  1.0 10.40  0.18
 65 NZ  LYS    4  1.0 10.40  0.02
 74 NZ  LYS    5  1.0 10.40  0.48
118 NZ  LYS   11  1.0 10.40  0.08
135 NH1 ARG   13  0.0 12.00  0.67
136 NH2 ARG   13  1.0 12.00  0.26
142 SG  CYS   14 -1.0  9.00  0.96
165 SG  CYS   17 -1.0  9.00  0.96
172 ND1 HIS   18  0.0  6.30  0.96
→175 NE2 HIS   18  1.0  6.30  0.96
197 OE1 GLU   21  0.0  4.50  0.63
198 OE2 GLU   21 -1.0  4.50  0.50
207 NZ  LYS   22  1.0 10.40  0.02
229 ND1 HIS   26  0.0  6.30  0.96
232 NE2 HIS   26  1.0  6.30  0.84
241 NZ  LYS   27  1.0 10.40  0.64
282 ND1 HIS   33  0.0  6.30  0.96
285 NE2 HIS   33  1.0  6.30  0.87
322 NH1 ARG   38  0.0 12.00  0.99
323 NH2 ARG   38  1.0 12.00  0.55
330 ND1 HIS   39  0.0  6.30  0.41
333 NE2 HIS   39  1.0  6.30  0.39
365 OE1 GLU   44  0.0  4.50  0.10
366 OE2 GLU   44 -1.0  4.50  0.02
382 OH  TYR   46 -1.0 10.00  0.98
400 OH  TYR   48 -1.0 10.00  0.96
414 OD1 ASP   50 -1.0  4.00  0.10
415 OD2 ASP   50  0.0  4.00  0.27
445 NZ  LYS   54  1.0 10.40  0.02
454 NZ  LYS   55  1.0 10.40  0.57
498 OD1 ASP   60  0.0  4.00  0.96
499 OD2 ASP   60 -1.0  4.00  0.54
507 OE1 GLU   61  0.0  4.50  0.88
508 OE2 GLU   61 -1.0  4.50  0.45
546 OE1 GLU   66  0.0  4.50  0.79
547 OE2 GLU   66 -1.0  4.50  0.09
559 OH  TYR   67 -1.0 10.00  0.96
598 NZ  LYS   72  1.0 10.40  0.26
607 NZ  LYS   73  1.0 10.40  0.02
619 OH  TYR   74 -1.0 10.00  0.67
654 NZ  LYS   79  1.0 10.40  0.59
703 NZ  LYS   86  1.0 10.40  0.81
712 NZ  LYS   87  1.0 10.40  0.20
720 OE1 GLU   88 -1.0  4.50  0.15
721 OE2 GLU   88  0.0  4.50  0.61
730 NZ  LYS   89  1.0 10.40  0.02
737 OD1 ASP   90 -1.0  4.00  0.64
738 OD2 ASP   90  0.0  4.00  0.80
748 NH1 ARG   91  1.0 12.00  0.57
749 NH2 ARG   91  0.0 12.00  0.97
764 OD1 ASP   93 -1.0  4.00  0.83
765 OD2 ASP   93  0.0  4.00  0.88
800 OH  TYR   97 -1.0 10.00  0.56
817 NZ  LYS   99  1.0 10.40  0.79
826 NZ  LYS  100  1.0 10.40  0.02
837 SG  CYS  102 -1.0  9.00  0.96
845 OE1 GLU  103 -1.0  4.50  0.10
846 OE2 GLU  103  0.0  4.50  0.62
847 OXT GLU  103 -1.0  3.50  0.54
848 FE  HEM  104  1.0 13.00  0.96
862 O1A HEM  104  0.0  4.00  0.96
863 O2A HEM  104 -1.0  4.00  0.96
889 O1D HEM  104  0.0  4.00  0.96
890 O2D HEM  104 -1.0  4.00  0.96

```

We use the editor to zero out the 1.0 charge on HIS18 NE2. Then we **save/exit** sites.list window and go back to the TK Menu.
Press **RETURN** to continue.

Now we are finally prepared to do a titration, or Tanford-Kirkwood calculation.

TITRATE

This command generates the following menu needed to set titration parameters:

```
-----Titration parameters-----
# description                      current value
-----
1  temperature (k)                  298.15
2  solvent dielectric constant      78.3
3  protein interior dielectric constant  4.0
4  solvent ionic strength           0.1000
5  protein effective radius         16.80
6  ion exclusion radius             18.80
7  mode of operation    0=a single pH    0
                          1=range of pH
8  pH to study (or lower value in range)  7.00
9  upper pH in range (when applicable)  10.00
10 pH increment (when applicable)      0.50
-----
input parameter # to change, 0 to run, -1 to cancel
```

At this point we could change system environmental variables such as temperature, solvent dielectric, interior dielectric, ionic strength of the medium, the pH, etc. If we changed the temperature, the solvent dielectric would automatically adjust to water's dielectric at this new temperature. To override this, one would need to re-enter the original dielectric value after resetting the temperature.

We want to work at an ionic strength of 0.15 so we will enter this. Type **'4'** and then input the new ionic strength.

0.15

Now run a titration by typing **'0'**.

At the finish of the TK titration calculation, the following window named TK_RESULTS appears:

```

*****
      Tanford-Kirkwood Calculation
      with Static Accessibility Modification
*****v3.0.0*****
Molecule = lycc.mdx

      Radius                16.80
Angstroms
      Exclusion radius      18.80
Angstroms
      charge burial depth   0.00
Angstroms
      exterior dielectric   78.305
      interior dielectric   4.000
      Temperature          298.15 K

pH= 7.0000  Ionic strength= 0.1500  T = 298.15 K

      Total stabilization energy = -2.4415 kcal/mol
      Number of iterations = 4
      Total charge= 7.642
      net ion= 0.000
      sum of delta pK = 3.5904

AA Atom  Res  Seg pKint  dpK    pK    Charge
THR  N     -5   8.00  0.138  8.138  0.932
GLU  OE1   -4   4.50 -0.017  4.483 -0.997
GLU  OE2   -4   4.50  0.000  4.500  0.000
LYS  NZ    -2  10.40  0.017 10.417  1.000
LYS  NZ     4  10.40 -0.008 10.392  1.000
LYS  NZ     5  10.40  0.140 10.540  1.000
LYS  NZ    11  10.40 -0.027 10.373  1.000
ARG  NH1   13  12.00  0.000 12.000  0.000
ARG  NH2   13  12.00 -0.110 11.890  1.000
CYS  SG    14   9.00 -0.530  8.470 -0.033
CYS  SG    17   9.00 -0.638  8.362 -0.042
HIS  ND1   18   6.30  0.000  6.300  0.000

HIS  NE2   18   6.30  0.000  6.300  0.000
GLU  OE1   21   4.50  0.000  4.500  0.000
GLU  OE2   21   4.50 -0.286  4.214 -0.998
LYS  NZ    22  10.40 -0.060 10.340  1.000
HIS  ND1   26   6.30  0.000  6.300  0.000
HIS  NE2   26   6.30 -0.078  6.222  0.143
LYS  NZ    27  10.40  0.041 10.441  1.000
HIS  ND1   33   6.30  0.000  6.300  0.000
HIS  NE2   33   6.30  0.623  6.923  0.456
ARG  NH1   38  12.00  0.000 12.000  0.000
ARG  NH2   38  12.00  0.126 12.126  1.000
HIS  ND1   39   6.30  0.000  6.300  0.000
HIS  NE2   39   6.30  0.078  6.378  0.193
GLU  OE1   44   4.50  0.000  4.500  0.000
GLU  OE2   44   4.50 -0.039  4.461 -0.997
TYR  OH    46  10.00 -0.684  9.316 -0.005
TYR  OH    48  10.00  0.363 10.363  0.000
ASP  OD1   50   4.00 -0.006  3.994 -0.999
ASP  OD2   50   4.00  0.000  4.000  0.000
LYS  NZ    54  10.40  0.038 10.438  1.000
LYS  NZ    55  10.40  0.094 10.494  1.000
ASP  OD1   60   4.00  0.000  4.000  0.000
ASP  OD2   60   4.00  0.049  4.049 -0.999
GLU  OE1   61   4.50  0.000  4.500  0.000
GLU  OE2   61   4.50 -0.160  4.340 -0.998
GLU  OE1   66   4.50  0.000  4.500  0.000
GLU  OE2   66   4.50 -0.054  4.446 -0.997
TYR  OH    67  10.00 -0.249  9.751 -0.002
LYS  NZ    72  10.40 -0.046 10.354  1.000
LYS  NZ    73  10.40 -0.077 10.323  1.000
TYR  OH    74  10.00 -0.248  9.752 -0.002
LYS  NZ    79  10.40  0.237 10.637  1.000
LYS  NZ    86  10.40 -0.259 10.141  0.999
LYS  NZ    87  10.40  0.038 10.438  1.000
GLU  OE1   88   4.50 -0.204  4.296 -0.998
GLU  OE2   88   4.50  0.000  4.500  0.000
LYS  NZ    89  10.40  0.066 10.466  1.000
ASP  OD1   90   4.00 -0.588  3.412 -1.000
ASP  OD2   90   4.00  0.000  4.000  0.000
ARG  NH1   91  12.00  0.154 12.154  1.000

```

ARG	NH2	91	12.00	0.000	12.000	0.000	GLU	OE2	103	4.50	0.000	4.500	0.000
ASP	OD1	93	4.00	-0.415	3.585	-1.000	GLU	OXT	103	3.50	-0.392	3.108	-1.000
ASP	OD2	93	4.00	0.000	4.000	0.000	HEM	FE	104	13.00	0.039	13.039	1.000
TYR	OH	97	10.00	-0.149	9.851	-0.001	HEM	O1A	104	4.00	0.000	4.000	0.000
LYS	NZ	99	10.40	0.052	10.452	1.000	HEM	O2A	104	4.00	-0.258	3.742	-0.999
LYS	NZ	100	10.40	-0.157	10.243	0.999	HEM	O1D	104	4.00	0.000	4.000	0.000
CYS	SG	102	9.00	-0.087	8.913	-0.012	HEM	O2D	104	4.00	-0.492	3.508	-1.000
GLU	OE1	103	4.50	0.034	4.534	-0.997							

The last column are the net effective charges that TK theory recommends for each titratable residue. The pK column gives the shifted pK of a site in its protein environment under the said conditions. These results appearing in the TK_RESULTS file can be saved under a different file name by using the SAVE AS feature of your system editor. After exiting this file out, go back to the text window where you will respond to the:

Hit RETURN to continue

Now when you have gone back to the TK Menu, you are ready to return to the Main Menu by typing

RETURN

At this point it would be a good idea to save the results you have just generated. The TK-recommended charges, the surface exposures of the various sites, etc will then be in the MDX file on disk:

SAVE

1.5 Assigning charges to the atoms using ASSIGN and ADJUST commands.

The TK-calculated charges are not yet really assigned to the atoms. They are in the MDX file in the column with variable name ACHG. The charge assignments for the molecule are made in the CHARGE variable column. It is the latter that are used in subsequent electrostatic field calculations and BD simulations. That being the case, at this point one would wish to assign charges in the lycc.mdx file according to the newly found TK estimates. To do this, use the ASSING command:

ASSIGN

which generates the following menu:

```

-----charge ASSIGNment menu-----
| EXIT - return to Main Menu          |
| SRC - assign standard residue partial charges to each atom |
| FORMAL - assign charges to formally-charged atoms only |
| TKRC - assign selectively from TK recommended charges |
| NEWTABLE - designate a new charge_table filename; |
|   Currently designated charge_table = /disk06/macrodox3.0.0/program/macprep.chg
-----v3.0.0-----

```

Respond using the TKRC option.

TKRC

You will then be asked the question:

```

Are you using partial or formal charges on the molecule?
  respond   partial
            formal
            cancel

```

How you answer this depends on how robust you desire your charge model to be. If you are going to eventually use this MDX file to do rigorous BD calculations and this molecule will be the central or target molecule (Protein I in the BD scheme), you would want to respond by saying 'partial'. If this is to be Protein II in that scheme, you would want to cut down the number of charges translating and rotating in the field of Protein I, and you should use 'formal' charges. Here let's answer

partial

A status list of charge assignments appears. Notice that the column of charges on the right are the TK suggested residue charges, which are also now in the ACHG(i) column of the MDX file. The middle column of charges are the currently assigned values, which are in the CHARGE(i) column of the MDX file. The ASSIGN function's job is to take, by user control, the ACHG charges from the TK calculation and selectively assign them into the CHARGE column.

Current charge assignments				4	LYS	1.000	1.000 NZ
Residue		Curr Assignments	TK-recommended	5	LYS	1.000	1.000 NZ
-5	THR	0.000	0.932 N	11	LYS	1.000	1.000 NZ
-4	GLU	-1.000	-0.997 OE1	13	ARG	1.000	1.000 NH2
-2	LYS	1.000	1.000 NZ	14	CYS	0.000	-0.033 SG
				17	CYS	0.000	-0.042 SG

18	HIS	0.000	0.000						
21	GLU	-1.000	-0.998	OE2					
22	LYS	1.000	1.000	NZ					
26	HIS	0.000	0.143	NE2					
27	LYS	1.000	1.000	NZ					
33	HIS	0.000	0.456	NE2					
38	ARG	1.000	1.000	NH2					
39	HIS	0.000	0.193	NE2					
44	GLU	-1.000	-0.997	OE2					
46	TYR	0.000	-0.005	OH					
48	TYR	0.000	0.000	OH					
50	ASP	-1.000	-0.999	OD1					
54	LYS	1.000	1.000	NZ					
55	LYS	1.000	1.000	NZ					
60	ASP	-1.000	-0.999	OD2					
61	GLU	-1.000	-0.998	OE2					
66	GLU	-1.000	-0.997	OE2					
67	TYR	0.000	-0.002	OH					
72	LYS	1.000	1.000	NZ					
73	LYS	1.000	1.000	NZ					
74	TYR	0.000	-0.002	OH					
79	LYS	1.000	1.000	NZ					
86	LYS	1.000	0.999	NZ					
87	LYS	1.000	1.000	NZ					
88	GLU	-1.000	-0.998	OE1					
89	LYS	1.000	1.000	NZ					
90	ASP	-1.000	-1.000	OD1					
91	ARG	1.000	1.000	NH1					
93	ASP	-1.000	-1.000	OD1					
97	TYR	0.000	-0.001	OH					
99	LYS	1.000	1.000	NZ					
100	LYS	1.000	0.999	NZ					
102	CYS	0.000	-0.012	SG					
103	GLU	-2.000	-1.996	OE1	OXT				
104	HEM	-2.000	-0.999	FE	O2A	O2D			

Total charge = 5.000 Tanford-Kirkwood tot charge = 7.642

Exit the Charge_Status_List file.

Hit RETURN to continue

In the text window will appear the menu of assignment commands:

```
--Charge assignments using Tanford-Kirkwood results--
EXIT; return to main menu
ALL; adjust every residue to tk values (exc hetero)
ALL <res name>; adjust every residue by that name
NUMBER ; adjust residue # to be specified
OXT; assign TK charge of atom OXT to oxy-terminal atom
NTE; assign TK charge of atom NT to N-terminal atom
EXACT; assign formal charges exactly as TK table
REDISPLAY; refresh current charge assignments table
```

give command:

First notice that TK recommends a 0.934 charge on the amino terminal Nitrogen. Responding with

NTE

makes this assignment and echos back a result:

```
N-terminal atom assigned a charge of  0.8521326
Done
give command:
```

Why does this say 0.852 rather than 0.934? It is because you are using a "partial charge" scheme for assigning charges, which means the formal charges of the residues, are all that TK calculation cares about, are to be smeared over the whole residue roughly according to a quantum of charge on the atoms. The N-terminal atom ends up with an 0.852 charge, but the net charge of that terminal residue will be 0.934.

Also, we would like our HIS residues to inherit the TK charge recommendations. Use:

ALL HIS

Now you may generate a new CHarge_Status_List window with these new assignments by using the redisplay command

RED

When the Charge_Status_List reappears note these changes have been accomplished, but still the Total charge = 6.723, whereas the total TK charge is 7.642. What's missing? Well, we haven't allowed the FE atom to inherit it's +1.0 charge, such that the TK value for HEM is -1.000 whereas the currently assigned net charge on HEM is -2.000. To treat the FE cannot be done here in this function. We must instead EXIT and go use the

ADJUST command in the CHARGES functions. First, you may close out the Charge_Status_List windows that appeared. Then:

Exit the Charge_Status_List

EXIT to return to main menu.

This should put you back to the Main Menu. Now enter the ADJUST function:

ADJUST

```
-----Adjust menu-----
1  Increment N-terminal charge;  now =   0.854
2  Increment Oxy-terminal charge; now =  -1.000

--Adjust another atomic charge,
   (atom selected by which format below?)
3 = atom name, charge assignment, mass(optional)
4 = res #, atom name, charge assignment, mass(optional)
5 = atom seq #, charge assignment, mass(optional)

-----
Enter parameter # for desired operation; 0 to RETURN
```

This menu is used to 'tweak' certain special atoms such as FE or other metal ions. Choose selection

3

which will allow us to reference that atom FE we wish to adjust by its atom name, which is uniquely 'FE '. We will now be asked the question:

Assign(0) or increment(1) existing charge?

We are in the partial charge scheme, so we want to increment the FE partial charge, which is right now by default equal to 0.24 in a reduced Fe-porphyrin(IV). We want to increment an amount of +1.0, so choose the increment option

1

for increment mode

```

name      charge      mass
.....
FE      1.000

```

FE Atom found. Charge is now = 1.240

NOTE: you must use spaces between FE and 1.000 and not a tab because MacroDox won't recognize the tab and think your atom doesn't exist!

Now enter 0 to return to the Main Menu.

At this point we should save what we have, so we don't lose any of the charge assignment work we have done.

SAVE

Now let's look at the status table to see if we are finished assigning charges:

STATUS

Current charge assignments									
Residue		Curr Assignments	TK-recommended						
-5	THR	0.932	0.932	N	60	ASP	-1.000	-0.999	OD2
-4	GLU	-1.000	-0.997	OE1	61	GLU	-1.000	-0.998	OE2
-2	LYS	1.000	1.000	NZ	66	GLU	-1.000	-0.997	OE2
4	LYS	1.000	1.000	NZ	67	TYR	0.000	-0.002	OH
5	LYS	1.000	1.000	NZ	72	LYS	1.000	1.000	NZ
11	LYS	1.000	1.000	NZ	73	LYS	1.000	1.000	NZ
13	ARG	1.000	1.000	NH2	74	TYR	0.000	-0.002	OH
14	CYS	0.000	-0.033	SG	79	LYS	1.000	1.000	NZ
17	CYS	0.000	-0.042	SG	86	LYS	1.000	0.999	NZ
18	HIS	0.000	0.000		87	LYS	1.000	1.000	NZ
21	GLU	-1.000	-0.998	OE2	88	GLU	-1.000	-0.998	OE1
22	LYS	1.000	1.000	NZ	89	LYS	1.000	1.000	NZ
26	HIS	0.143	0.143	NE2	90	ASP	-1.000	-1.000	OD1
27	LYS	1.000	1.000	NZ	91	ARG	1.000	1.000	NH1
33	HIS	0.456	0.456	NE2	93	ASP	-1.000	-1.000	OD1
38	ARG	1.000	1.000	NH2	97	TYR	0.000	-0.001	OH
39	HIS	0.193	0.193	NE2	99	LYS	1.000	1.000	NZ
44	GLU	-1.000	-0.997	OE2	100	LYS	1.000	0.999	NZ
46	TYR	0.000	-0.005	OH	102	CYS	0.000	-0.012	SG
48	TYR	0.000	0.000	OH	103	GLU	-2.000	-1.996	OE1 OXT
50	ASP	-1.000	-0.999	OD1	104	HEM	-1.000	-0.999	FE O2A O2D
54	LYS	1.000	1.000	NZ					
55	LYS	1.000	1.000	NZ					

Total charge = 7.723 Tanford-Kirkwood tot charge = 7.642

The slight discrepancy between the total TK charge 7.642 and the total assigned charge 7.723 is due

to the accumulated effects of the ASP and GLU, which are not quite -1.000 in the TK calculation, but we will leave them as -1.000 in our work. You may **exit** down this window Charge_Status_List if desired.

Now let's see what the dipole moment is:

PROPERTIES

```
lycc.mdx
number of atoms=          890
Total mass= 12671.35 amu
Center of mass =   0.000   0.000   0.000
Center of +/- charge= -0.444   0.020  -0.641       0.641  -0.550  -0.430
Total +/- charge= 82.435  74.711 Net Charge=   7.723e
Dipole magnitude=  465.82 Debye
Dipole vec =-405.982 205.567 -99.534
Greatest distance from center of mass =  22.584 A
Full surface calculation not performed yet
Radius of gyration=  12.72
Atom # density in concentric shells of thickness=  3.0
  shell #  density
    1      35.4    2      41.7    3      54.0    4      62.6
    5      40.3    6      15.5    7      2.3    8      0.2
Estimated effective spherical radius =  16.8 A
Est Stokes-Einstein trans and rot diff coeff =
0.130E-01 A**2/ps, 0.345E-04 1/ps
Heme group found
Heme plane normal vector:  -0.628  -0.776  -0.051
Angle between dipole and heme normal =  77.53
Angle between dipole and Fe vector =  80.97
```

Exit PROPERTIES.LIS window

Generate listing of atoms along dipole vector axes?

yes

The response briefly shows and then disappears.

```
-405.9825      205.5671      -99.53417
Listing sent to file DIPOLE.LIS
```

Now we see a better value for the dipole moment = 465.82 Debye.

Let us also answer yes to the last question and generate a listing of atoms along the dipole vector axes.

The following listing gets sent to file DIPOLE.LIS:

atom name	residue	residue no.	angle	distance	CG	ARG	38	170.022	11.549
CE	LYS	5	27.699	18.202	CD	ARG	38	164.627	11.683
NZ	LYS	5	24.628	17.461	NE	ARG	38	161.611	10.496
CA	LEU	9	24.712	11.967	CZ	ARG	38	154.386	10.390
C	LEU	9	27.366	11.005	NH1	ARG	38	150.064	9.407
O	LEU	9	25.152	10.862	NH2	ARG	38	150.868	11.484
CB	LEU	9	22.880	11.288	N	HIS	39	168.734	12.044
CG	LEU	9	15.782	11.442	CA	HIS	39	165.645	11.873
CD1	LEU	9	14.983	12.924	C	HIS	39	168.081	11.047
CD2	LEU	9	16.525	10.837	O	HIS	39	173.532	10.950
C	THR	12	29.474	12.946	CB	HIS	39	164.073	13.272
O	THR	12	29.485	13.552	CG	HIS	39	161.075	14.206
CB	THR	12	25.790	14.769	ND1	HIS	39	162.881	15.008
OG1	THR	12	22.179	14.092	CD2	HIS	39	155.887	14.651
CG2	THR	12	27.480	16.019	CE1	HIS	39	158.996	15.783
N	ARG	13	28.221	11.666	NE2	HIS	39	155.267	15.660
CA	ARG	13	27.052	10.856	N	SER	40	161.805	10.666
CB	ARG	13	19.152	10.314	CA	SER	40	158.716	10.104
CG	ARG	13	14.430	11.524	C	SER	40	157.970	11.248
CD	ARG	13	14.774	12.310	O	SER	40	158.837	12.404
NE	ARG	13	11.708	11.984	CB	SER	40	150.536	9.832
CZ	ARG	13	15.559	12.585	N	GLY	41	155.435	10.943
NH1	ARG	13	16.300	12.577	CA	GLY	41	152.905	12.120
NH2	ARG	13	20.160	13.267	C	GLY	41	156.611	12.939
O	ILE	35	152.920	8.503	O	GLY	41	155.119	14.141
CG1	ILE	35	160.504	5.931	N	GLN	42	160.548	12.550
CD1	ILE	35	166.825	6.312	CA	GLN	42	162.637	13.570
C	GLY	37	152.222	12.391	C	GLN	42	159.306	13.620
O	GLY	37	153.881	13.501	O	GLN	42	159.428	14.545
N	ARG	38	156.208	11.486	CB	GLN	42	168.872	13.859
CA	ARG	38	163.467	11.806	CG	GLN	42	170.863	14.242
C	ARG	38	165.325	11.303	CD	GLN	42	169.040	15.554
O	ARG	38	160.827	10.453	OE1	GLN	42	166.312	15.743
CB	ARG	38	164.860	11.562	NE2	GLN	42	169.665	16.598

N	ALA	43	155.221	12.677	CG	LYS	87	27.988	17.032
CA	ALA	43	150.092	12.949	CD	LYS	87	23.949	17.781
CG	ASN	56	152.755	15.992	CE	LYS	87	23.114	19.029
OD1	ASN	56	152.989	17.072	NZ	LYS	87	25.041	18.997
ND2	ASN	56	156.959	15.389	NZ	LYS	89	29.022	20.417
O	VAL	57	152.023	10.857	CA	ASP	90	28.620	13.000
CD2	LEU	58	151.573	13.424	CB	ASP	90	23.743	13.077
CG	TRP	59	153.846	6.855	CG	ASP	90	20.767	14.367
CD1	TRP	59	163.472	7.260	OD1	ASP	90	18.477	14.841
NE1	TRP	59	160.238	6.963	OD2	ASP	90	21.804	15.030
CE2	TRP	59	150.218	6.348	CD2	LEU	94	21.875	6.316
CE2	PHE	82	28.063	9.612	CMA	HEM	104	153.273	2.445
CZ	PHE	82	25.358	8.917	CGA	HEM	104	153.728	7.254
O	GLY	84	29.967	12.002	O1A	HEM	104	155.759	7.885
CA	LEU	85	27.792	11.116	O2A	HEM	104	153.631	7.742
CB	LEU	85	24.210	10.011	CMB	HEM	104	28.498	3.183
CG	LEU	85	15.552	9.798	CAB	HEM	104	27.608	6.107
CD1	LEU	85	13.267	8.712	CBB	HEM	104	17.507	6.519
CD2	LEU	85	12.561	11.078					
CB	LYS	87	29.613	15.890					

The dipole vector passes within 12 degrees of ARG 13 atom NE.

Now let's get out of the program by saving and quitting

SAVE

QUIT