**File:hydromic7c.doc**

# HYDROMIC

Version 7c, January 2006

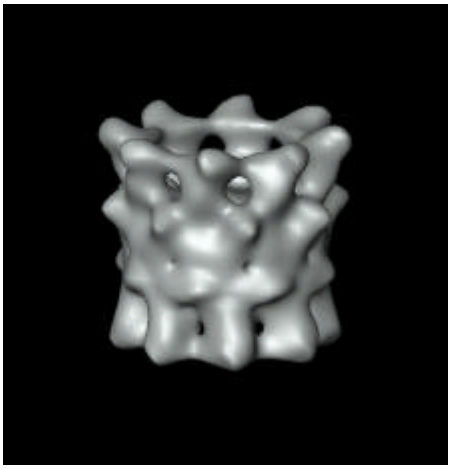
## Index

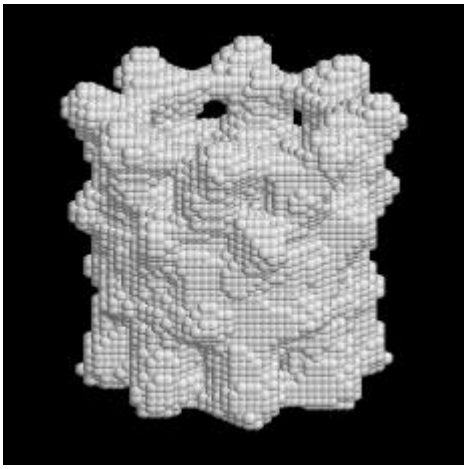


## 1. Introduction to HYDROMIC

HYDROMIC is a computer program for the calculation of solution properties of macromolecular or supramolecular structures determined by cryo-electron microscopy or related techniques. The structure is specified by a 3D density map or histogram, contained in a file which can be (unfortunately!) is a variety of formats. The so-called *spider* file is amongst the mos popular, and still there can be different spiders depending on the computer where it has been generated. Structures of large macromolecules, and macromolecular complexes may be available in such file. HYDROMIC reads the structural file, builds from it an appropriate hydrodynamic model, and proceeds with the calculation of properties, including hydrodynamic coefficients, scattering-related quantities, etc.

In the following figure we show the cryo-electron microscopy reconstruction of the structure of CCT chaperonin and its hydrodynamic model.





## 2. Literature

The primary reference for HYDROMIC is:

- J. García de la Torre, J.M. Valpuesta and J.L. Carrascosa, "HYDROMIC: Prediction of hydrodynamic properties of rigid macromolecular structures obtained from electron-microscopy images" *Eur. Biophys. J.*, 30, 457-462 (2001).

If you employ scattering related properties (distribution of distances, longest distance, scattering function), or the covolume, then the you may also cite the reference where these calculations are described:

- J. Garcia de la Torre, B. Carrasco and S. E. Harding, "Calculation of NMR relaxation, covolume and scattering-related properties of bead models using the SOLPRO computer program", *Eur. Biophy. J.*, 28, 119-132" (1999).

If you wish to cite also the theoretical work on which the bead modeling procedure is based, a proper cite is our 1981 review in *Quarterly Reviews of Biophysics*. In our 1999 paper in *Biophysical Journal*, you can find an update of the theory, and a discussion on bead and shell modeling methodologies:

- J. Garcia de la Torre and V.A. Bloomfield, "Hydrodynamic properties of complex, rigid, biological macromolecules. Theory and applications". *Q. Rev. Biophy.*, 14, 81-139 (1981)

- B. Carrasco and J. Garcia de la Torre, "Hydrodynamic properties of rigid particles. Comparison of different modeling and computational strategies". *Biophys. J.* 76, 3044-3057 (1999).

## 3. Running HYDROMIC. Input data files

As in all the HYDRO family of programs, you will have to supply two input data files: (a) the main input data file, which will specify primary data such as temperature, solvent density, etc.; and (b) a structural data file, which will contain the information about the structure or geometry of the macromolecule or particle that you are considering. This structural file contains basically a 3D density map, which can be in different formats (see below) The name and type of the structural file will be one of the data in the main input data file.

The name of the main input data file for HYDROMIC will be `hydromic.dat,` and it will contain the following lines:

**3.a. First part of the input file**.

Basically this part provides the information on the structure of the macromolecule. It contains the following lines (the FORTRAN types are specified):

- TITLE(`CHARACTER*20`) Title of the calculation

- FILENAME (`CHARACTER*30`)Name to be used for the various output files corresponding to each subcase in a many-cases execution. The various files produced will have names of the form `filename.xxx`, where xxx is an extension depending on the file type (see section 4 below). This name would eventually include the path for the files.

- INPUT (`CHARACTER*30`) Full name of the structural file, eventually including its extension and path.

- THRESOLD (`REAL`) Minimum density for a pixel to be considered in the particle

- SPACING (`REAL`) Spacing (angstroms) in the 3D grid. It is assumed to be the same in the three directions

- ITYPEMIC (`INTEGER`) Flag that indicates the format of the structural file.Valid values are:
  = 1 for *spider* files from SGI/SUN
  = 2 for *spider* files from DEC/LINUX
  = 3 for MRC files
  = 4 for FORMATTED files, converted from other formats using the **`em2em`** program (Very important: when **`em2em`** asks "Give format string" , give 6F10.4)

  - ONLY If ITYPE = 4, Three more lines with the values (that are supplied by **`em2em`** when converting other formats to FORMATTED) :
    1) Number of pixels per line
    2) Number of lines per image
    3) Number of sections

HYDROMIC employs the shell-model methodology. The primary hydrodynamic model is not used in the hydrodynamic calculations. Instead, a shell model , composed of 'minibeads' of radius sigma is derived from it, and extrapolation to the shell model limit of sigma=0 is carried out. For this purpose, the following information has to be supplied next in the main input data file

- NSIG (`INTEGER`) Number of values of the radius of the minibead. The radius will range from SIGMIN to SIGMAX. It must be greater than 2 (typically 5 to 8). There is also the possibility of letting to the program the task of estimating the two extreme values of sigma; this will be indicated giving the value –1 for NSIG, and in this case you will omit the two following lines with the values of SIGMIN and SIGMAX.

- SIGMIN (REAL) Lowest value of sigma, the minibead radius

- SIGMAX (REAL) Highest value of sigma, the minibead radius

The smaller the size of the minibeads, the larger the number of them needed to cover the surface of the particle. The present version of the program works with a maximum of 2000 minibeads. If SIGMIN is too small, an error message will be obtained and the program will stop. The value of SIGMAX should be taken such that the number of minibeads is not too small, say in the range 200-400.

**3.b. Second part of the input file**.

This part provides information on some basic properties of macromolecule and solvent. It contains the following lines:

- T (REAL) Temperature, Kelvin

- ETA (REAL) Solvent viscosity, poises

- RM (REAL) Molecular weight .

- VBAR (REAL) Partial specific volume, $cm^3/g$

- RHO (REAL) Solution (approx. Solvent) density, $g/cm^3$

**3.C. Third part of the input file**.

This part is intended for the calculation of some non-hydrodynamic properties that, in previous versions of our software, were considered separately in the computer program SOLPRO. These properties are the scattering form factor (Debye expression), the distribution of intramolecular distances, and the covolume, and are calculated from the coordinates of the spheres.

The data that you have to supply are:

- NH(INTEGER), the number of values of the scattering (angular) variable, h. If you wish to omit the scattering calculation, the value given here should be 0, and the following (HMAX) line is omitted.

- HMAX (REAL), the largest value of h ($cm^{-1}$), so that the scattering variable will range from 0 to HMAX

- NS(INTEGER), the number of intervals for the distribution of distances. The values of the intramolecular distances will be varied between 0 and the longest distance,

which is determined by the program. If you wish to omit the calculation of the distribution of distances, the value given here should be 0.

- RMAX (`REAL`), the maximum length for the calculation of the distribution of distances (in cm). If you give any negative value, it is understood that RMAX will be the longest distance in the model, which is calculated by the program. The value will be given in the same units as the coordinates and dimensions in the structural file

- NTRIALS (`INTEGER`) is the number of trials or MonteCarlo moves in the calculation of the covolume. Set this value to 0 if you wish to omit the covolume calculation.

In HYDROMIC, the calculation of scattering and distances is a bit lengthy, but still feasible. However, the calculation of the covolume can be extremely long. It may be advisable to avoid these calculations in the way indicated above.

The most important quantities related to translational and rotational diffusion are the translational and rotational diffusion coefficients and the rotational relaxation times. The program will give you these quantities. For some special purposes you may also want the full 6x6 diffusion tensor, which contains the 3x3 translational diffusion tensor, rotational diffusion tensor, and the translation-rotation coupling tensor, as well as the center of diffusion.

- IDIF is a flag that indicates (if IDIF is 1) that you wish a detailed report of the diffusivity of the particle, including the full (anisotropic) translational, rotational and coupling tensor, and the position of the hydrodynamic (diffusion) center.

**3.D. End of calculation or next case**

- Next or final line: If this case is the only or the final one, in the next line you will put an asterisk followed by 19 spaces.

## 4. Output files

Several files are produced at execution time. There will be a set of files for each case included in a single run. All these files will have a common name, given by the filename specified for each case in the input file, and a different extension. The extensions correspond to:

- `.res` Output file containing the mains results, with the name specified in the input file
- `-pri.bea` is a pdb-formatted file representing the primary hydrodynamic model, that can be viewed with RASMOL
- `-pri.vrml` is a VRML-formatted file containing the coordinates of primary hydrodynamic model, that can be viewed with a VRML viewer, like Cortona(for models with only subunits, without additional spheres). See our separate VisualBeads user Guide.

- `.sol` is an ASCII file containing data needed if you wish to run the separate program SOLPRO

Other files are:

- `summary.txt` is a numeric archive containing a line for each case in the calculation.

(a) If there was no distance distribution calculation and no scattering calculation, this file has 13 columns containing (1) first 10 characters of title; (2) translational diffusion coefficient; (3) radius of gyration; (4) volume; (5-9) the five relaxation times; (10) intrinsic viscosity; (11) sedimentation coefficient; (12) longest distance; (13) covolume. (b) If there was either distance distribution

(b) If there was scattering function calculation, *P(h)*, with a number of values NH, with NH not greater than 100, then the above columns will be followed by NH+2 columns containing NH, HMAX, and the values of the distribution function corresponding to the NS intervals (centered at $h_i$=[i+0.5]HMAX/NH

(c) If there was distance calculation *p(r)*, with a number of values NS, with NS not greater than 101, then the above columns will be followed by NS+2 columns containing NS, RMAX, and the values of the distribution function corresponding to the NS intervals (centered at $r_i$=[i+0.5]RMAX/NS

## 5. Hints and notes.

- The molecular weight is used in the calculation of the intrinsic viscosity and the sedimentation coefficient, and the solution density and the specific volume of the solute are used only for the sedimentation coefficient. If you do not know these quantities, you may give then some approximate or estimated values in the data files. HYDROMIC will still be useful, because all the other solution properties (diffusion coefficients, relaxation times, radius of gyration, scattering properties, covolume, etc) will be correct.

- The NSIG=-1 mode, for automatic estimate of SIGMAMIN and SIGMAMAX, may be slow, or not work at all, because the primary hydrodynamic model contains a large number of elements (as many as positive voxels).

- **`em2em`** is a simple and useful program for the conversion of formats of images from cryo-EM and other sources, kindly provided by Image Science. It can be downloaded from http://www.imagescience.de/em2em/index.htm

- **`IMPORTANT:`** Please read the HydroFAQs, frequently asked questions on the HYDRO programs for further information.

## 6. Release notes

This is the second released version of HYDROMIC. The previous version was 5a, of March 2002.

- The main changes is that the variety of formats that HYDROMIC can read has been greatly expanded. In the first version, only spider files from SGI/SUN (with the binary format specific of those platforms) could be processed. Now, the program admits also spiders with DEC/LINUX binary format, the widely used MRC format, and an ASCII. FORMATTED file that can be obtained from other formats using **`em2em`**.

- The number of pixels in the primary model may be up to 1 000 000 (it was 300 000) in the previous version

- Calculation of non-hydrodynamic properties: radius of gyration and volume of the model, scattering function, distribution of distances and covolume

- File `summary.txt` contains the results for the distance distribution and the scattering function.

- Visualization of the model in the VRML format

This program has been developed in a Windows PC. The MS DOS/Windows executable can be started from Windows, but we advise to open a MS DOS session in a window for program execution, while doing the other tasks (editing, visualization, etc) as usually in Windows.