File:hydro++10_54-PUB-Sept-2011.doc

HYDRO++

Version 10 , September 2011

(compilation 54)

Index

1. Introduction to HYDRO

HYDRO is a program for the calculation of hydrodynamic coefficients and other solution properties of rigid macromolecules, colloidal particles, etc, employing bead models. In strict sense, bead models are those in which the shape and volume of the particle are modeled as an array of spheres (beads) of arbitrary – equal or unequal – size. It is the classical program that has been available from us for many years. Note that there is another, different, modeling strategy based on bead-shell models. This alternative methodology is implemented in other programs of our suite (HYDROPRO, HYDRONMR, HYDROMIC, HYDROPIX, HYDROSUB, etc). The user of HYDRO has to build previously the bead model, which will be specified as a list of Cartesian coordinates and radii of the beads.

HYDRO++ is a revised and improved version of the well-known **HYDRO** program, which solves definitively some ambiguities related to rotational diffusion and the intrinsic viscosity. These properties are now evaluated with great accuracy. Results for translational (diffusion and sedimentation) coefficients and other solution properties are practically the same as with the **HYDRO** program.

The newest, version 10, is externally very much the same as the previous version 8, and contains minor fixes respect to the preliminary released version 9beta. However, internally is a completely new code, written now in Fortran 90, in a notably modular form. This has been done thinking of future extensions of the code, which eventually would become Open Source. At the moment, the re-programming has resulted in a great improvement in efficiency (calculation speed). This is not only because of the most efficient executables produced by the Fortran 90 compilers. Additionally, we have

introduced the possibility of parallel calculation, feasible in today's dual- or quad- core processors, or multi-processor computers. All this brings an extraordinary speed up improvement.

2. Literature

The newest reference for **HYDRO**, which describes the advances in **HYDRO++**, is:

- J. García de la Torre, G. del Rio and A. Ortega, "Improved calculation of rotational diffusion and intrinsic viscosity of bead models for macromolecules and nanoparticles", *J. Phys. Chem. B* 111, 955-961 (2007).

which is based on the original, classical publication:

- J. Garcia de la Torre, S. Navarro, M.C. Lopez Martinez, F.G. Diaz, J. Lopez Cascales. HYDRO. A computer software for the prediction of hydrodynamic properties of macromolecules. *Biophys. J.* 67, 530-531 (1994).

If you employ scattering related properties (distribution of distances, longest distance, scattering function), or the covolume – which is related to the second virial coefficient –then 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. Biophys. J.*, 28, 119-132 (1999).

You may also wish to cite 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". *Biophysical Journal* 76, 3044-3057 (1999).

3. Running HYDRO++. Input data files

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.

You may launch **HYDRO++** clicking the icon in MS/Windows or – what is better recommended - at the system prompt in a MS/DOS console. You can do it similarly at the system prompt in a Linux console. Then, the program will ask you for the name of the input data file. Alternatively, you can write batch (script) files to launch the program automatically containing the instruction to execute the program and the name of the main input file. The name of the structural file will be one of the data in the main input data file (see below).

An unlimited number of cases can be executed in a single run of **HYDRO++**, with a single main input file, which would contain a series of blocks of data, one for each case. Each block will contain the following lines:

3.a. First part of a block of data

Basically this part provides the information on the structure of the macromolecule or particle that is being modelled. 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)` Name of the (separate) file containing the information on the bead model, eventually including its path.It will have the following lines: (1) `(REAL)` U, unit of length (in cm) in which coordinates and radii are given (2) `(INTEGER)` Number of beads, and (following N lines), in each line, four comma-or space- separated values: three Cartesian coordinates and radii of beads.

- ICASE `(INTEGER)`, type of hydrodynamic calculation. Valid cases are:

  ICASE = 12 Kirkwood-Riseman calculation with modified Oseen tensor, with full volume correction for radius of gyration and rotational properties, and adjusted volume correction for the intrinsic viscosity. This option – new in **HYDRO++** - is to be preferred to the previous one.

  ICASE = 20 Third-order hydrodynamic interaction that eliminates the need of volume correction for rotation. The intrinsic viscosity is not calculated in this case. Computing time is slightly longer than for ICASE=12

  ICASE = 21 Cubic substitution, that eliminates the need of volume correction for rotation and intrinsic viscosity and is most accurate for all the properties. Computing

time is appreciably longer, but this is unimportant if the number of beads is not very large

3.b. Second part of a block of data .

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

- T (REAL) Temperature, centigrade
- 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 a block of data.

This part is intended for the calculation of some non-hydrodynamic properties, namely the scattering form factor (angular dependence of scattering intensities), 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:

- NQ(INTEGER), the number of values of the scattering (angular) variable, q . If you wish to omit the scattering calculation, the value given here should be 0 (Recommended values are multiples of 5 or 10 plus 1, i.e., 11, 51, etc; value no. 1 will be q=0 with S(q)=1). There is the option of an automatic choice of the scattering data, NQ and QMAX, which is activated giving -1 for NQ.
- QMAX (REAL), the largest value of q (cm^{-1}), so that the scattering variable will range from 0 to QMAX, This line will be omitted if there is no scattering calculation (NQ =0), of if it is done with automatic values (NQ =-1).
- 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. Also, there is the possibility of an automatic choice of data for this calculation, which is indicated giving -1 for NS

- RMAX (REAL), the maximum length (in cm) for the calculation of the distribution of distances (in cm). This line will be omitted if there is no distribution of distance calculation (NQ =0), of if it is done with automatic values (NQ =-1).

- 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. Recall that this calculation is very time-consuming, and must be restricted to models with not too many beads.

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 and 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 name derived from the `filename` specified for each case in the input file, and a different extension. The extensions correspond to:

- `-res.txt` Output file containing the mains results, with the name specified in the input file

- `.bea` is a a pdb-formatted file containing the coordinates of the original bead model (mode HYDRO), to be viewed with any pdb viewer, such as RASMOL. This is intended for models with identical beads. If all the beads are not of the same size, the radius taken for all of them will be that of the first bead. NOTE: Do not forget to specify: Display / SpaceFill in the RasWin menu to see the visualization files.
- `.vrml` is a a VRML-formatted graphics file, which accepts beads of unequal sizes. This is to be handled using a VRML viewer, such as CORTONA's. (See our VISUALBEADS software for further information).
- `-sol.txt` is an ASCII file containing data needed if you wish to run the separate program SOLPRO

Other files are:

- A "summary" file whose name will be derived from the name of the main input file, trimming the last four characters of that name and adding "-sum.txt" (for instance if the main input file was mydata.dat, the summary file will be mydata-sum.txt). This file 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

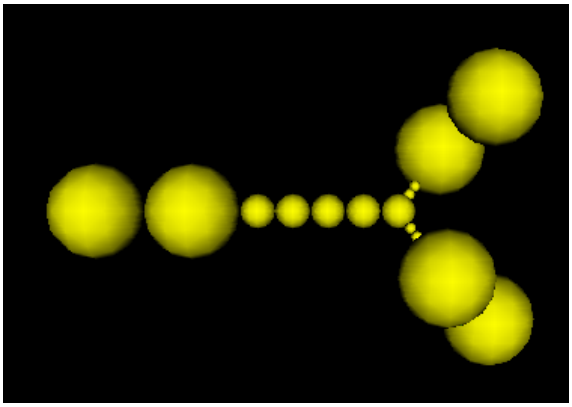
(b) If there was scattering function calculation, *S(q)*, with a number of values NQ, with NQ not greater than 100, then the above columns will be followed by NQ+2 columns containing NQ, QMAX, and the values of the distribution function corresponding to the NQ values.

(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)

- A file intended for the use of the results in the analysis of experimental data using the HYDFIT program (see the User Guide of that program for more details). This file will be named with the same procedure as the summary file, in this case ending with "-fit.txt" (in the above example, mydata-fit.txt).

5. Examples.

As a first example, we present results for a simple bead model of the human antibody molecule IgG3 (Davis et al, *Molecular Immunology*, **1987**, 24, 821-829). This model, which has only N=15 beads, has been traditionally posted in our web as the example of the previous versions of HYDRO.

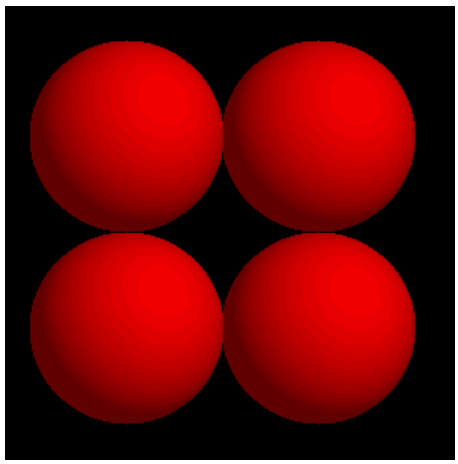


Results for the main solution properties are given below:

	ICASE	D_t x 10^7 cm^2/s	τ_h x 10^7 s	$[\eta]$ cm^3/g
Adjusted volume correction for $[\eta]$	12	3.76	2.80	8.71

3rd order HI procedure	20	3.70	2.50	8.71
Cubic substitution	21	3.63	2.57	9.20

Another example is a square tetramer, having four touching beads:



The following table contains the results for the various approaches. Along with the results from **HYDRO++** with the various approaches, we include results (which are practically exact) of a shell-model calculation with our **HYDROSUB** program. As for rotational properties, we report in this case the two distinct components of the rotational diffusion tensors, corresponding to axes in the plane of centers (para) and perpendicular to the plane (perp). We also report the CPU time required for the calculation (for more information on CPU times, see next section.

	ICASE	D_t x 10^7 cm^2/s	τ_h x 10^7 s	$D_r^{(para)}$ x 10^{-6} s^{-1}	$D_r^{(perp)}$ x 10^{-6} s^{-1}	$[\eta]$ cm^3/g	CPU time (approx), seconds
Adjusted volume correction for $[\eta]$	12	6.26	7.37	1.89	2.45	3.99	0.8 x 10^{-4}
3rd order HI procedure	20	6.01	5.42	2.57	3.32	3.99	1.0 x 10^{-4}
Cubic substitution	21	5.84	5.66	2.41	3.21	3.88	1.2 x 10^{-3}
SHELL MODEL (practically exact)	HYDRO-SUB	5.84	5.60	2.41	3.23	3.95	300

It is clear that the 3rd order and cubic-substitution method give practically exact results with a CPU time several orders of magnitude shorter.

(Note: CPU time is for a single-core execution. Information on multi-core CPU timing is given below).

6. Hints and notes.

- Temperature is given in degrees Celsius (centrigrade)

- We recall the possibility, indicated in section 3, of running an unlimited cases in a single execution, with a single main input file containing a series of blocks, one for each case

- Dynamic memory allocation was employed in programming in order to make the number of beads not limited by the code. However, the limit of accessible memory in each computer platforms, depending on the operating system and the amount of memory in the computer, puts a runtime limit in the number of beads. For a Windows computer with 4 GB, the limit is around 6000. Linux allows even larger limits.

- The working modes ICASE=-1 (automatic choice) and 11 (full volume correction) have disappeared in this version.

- Regarding the number of beads and the working mode, the choices are left to the user (obviously, more beads means larger particle or more detail, and the working mode can affect – although quite slightly to the precision of the results). It is advisable to make several trials…

- The molecular weight is used in the calculation of the intrinsic viscosity and the sedimentation coefficient, and the specific volume and solution density are only used for the sedimentation coefficient. If you do not know these quantities, you may give then some approximate or estimated values in the data files. HYDRO++ will still be useful, because all the other solution properties (diffusion coefficients, relaxation times, radius of gyration, scattering properties, covolume, etc) will be correct.
- As indicated above, the cubic substitution (ICASE=20) gives the most accurate results for all the properties. If your structures have a moderate number of beads, or if you have to run a few structures, this would be the method of choice. If you do not need the intrinsic viscosity, the 3rd order procedure (ICASE=21) gives very reliable results for the rotational quantities.

- HYDRO++ is intended for models with non-overlapping spheres. If it were applied to modes with overlapping spheres, the results for the rotational properties and the intrinsic viscosity may be really wrong. If you want to build your model with overlapping spheres, the hydrodynamic calculations can be made with our program HYDROSUB.

- 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 (rather than directly clicking on the icon of the exe file), while doing the other tasks (editing, visualization, etc) as usually in Windows. An executable of HYDRO is also available for Linux.

- In this version there are some minor modifications in the adjusted volume-correction for the intrinsic viscosity, as described in J. García de la Torre, D. Amorós and A. Ortega, "Intrinsic viscosity of bead models for macromolecules and nanoparticles", *European Biophysics Journal* 39, 381-388 (2010). This may result in a slightly different (better) results for the intrinsic viscosity with those of previous versions of HYDRO

- For more information, see the HydroFAQs (frequently asked questions about the HYDRO programs)

- Now, among the results in the output file, you will find a list of the equivalent radii corresponding to the various properties. For a description of the definition and utility of the equivalent radii, see: A. Ortega, J. García de la Torre, "Equivalent radii and ratios of radii from solution properties as indicators of macromolecular shape, conformation and flexibility" *Biomacromolecules* 8, 2464-2475 (2007)

6. Release notes

As indicated above, version 9beta was fully reprogrammed in Fortran 90. However, for users, changes are minimal:

- Temperature is given in degrees Celsius (centrigrade)

- Working modeS ICASE=-1 (automatic choices) 10 (no volume corrections) and 11 (full volume corrections) have been suppressed.

- Note that there is new option of "automatic" data for the calculation of scattering form factors and distribution of distances

- Summary and HYDFIT files are now named differently