1 Manual Cover Page

SEDNTERP 3 User Manual:

Theory, Equations, and References

Program Elements, Forms, and Features

Update For Program Version 3.04

January 17, 2025

Table of Contents

2 Introduction

2.1 Overview of SEDNTERP

What is the purpose of SEDNTERP?

The SEDNTERP program is a general purpose tool for the interpretation of sedimentation equilibrium, sedimentation velocity, and diffusion experiments. It collects many diverse functions into a single program. SEDNTERP is useful both for computing molecular structural estimates from experimental data and for interpolation of physical parameters of common buffers used in sedimentation experiments. It is also useful to create a database of:

- the properties of samples (proteins) you work with (molar mass, partial specific volume, hydration, UV extinction coefficients, *dn/dc*)
- the properties of buffers that you use in your lab (density, viscosity, refractive index, pH, composition)
- experimental results (sedimentation coefficients, diffusion coefficients, sedimentation equilibrium results)
- the computed hydrodynamic properties of your samples (maximum possible sedimentation and diffusion coefficients, f/f_0 ratio, Perrin P function, computed shape and dimensions from ellipsoidal or cylindrical models)

Importantly, storing this information in the program database allows you to document and reproduce any calculated values that you will be using in reports or publications.

What is its rationale, and what approach is used in doing the calculations?

Studies of the structures of macromolecules and their complexes are often aided by thermodynamic and hydrodynamic data available from analytical ultracentrifugation. These techniques are amongst the most versatile, rigorous and accurate for estimating useful physical parameters for a wide range of chemical systems. The equations describing sedimentation are derived from first principles (Ref. 1, Ref. 2, Ref. 3). All terms in the mathematical equations that describe sedimentation behavior are either readily and accurately estimable (Ref. 4, Ref. 5) or experimentally determinable (Ref. 6, Ref. 7). The quantity and quality of the information furnished is great given the relatively small effort needed to perform the experiments. The sheer number of calculations and their inter-relations, however, can seem overwhelming to the non-expert. For the full power of analytical sedimentation to be realized it is necessary for many of the routine calculations to be automated.

The use of appropriate polynomial functions to represent (interpolate) tabulated data permits the construction of a program that uses readily available experimental parameters (temperature, buffer composition, *etc*.) in the computations needed for the interpretation of analytical sedimentation data. Proper design of the program permits the maximum interpretation of data while requiring only a minimum number of experimental observations. At the same time, any calculations which cannot be made with certainty should not be performed, thus preventing over-interpretation of data. The program is designed so that user-supplied information always takes precedence in calculations. Thus, as more experimental data is available, more rigorous interpretations may be made.

SEDNTERP stores the basic physical data and the fitted coefficients for interpolations in its database file. Storing all of this information in a single file helps in the program set up and in the updating of the algorithms. Likewise, information entered into the program by the user can (and should) also be saved into the program database allowing users to document and reproduce the calculations they have made, and also to save time by avoiding the need to re-enter information about the proteins and buffers they commonly work with in their laboratory.

The program is designed to be very comprehensive and flexible in using all possible forms of information available from experiment. However, SEDNTERP will **assume** default parameters for quick analysis. Therefore, it is the responsibility of the user to supply SEDNTERP with all available experimental information in order to get the most accurate interpretation. Particularly when many default values are used, the accuracy of the computations performed by SEDNTERP will be questionable.

All computations are carried out using 15-digit ("double") precision.

The executable files are free and may be copied and distributed freely. There is no guarantee or warranty implied about the accuracy of this program. The program source code is also available to interested scientists. Comments, problems, and suggestions are sincerely requested from users.

A brief history of the program

SEDNTERP was originally developed as a template for a spreadsheet by Thomas M. Laue, Bhairavi D. Shah, Theresa M. Ridgeway and Sandra L. Pelletier. Their work is described in Ref. 59.

Dr. David P. Hayes (david.hayes@sednterp.org, now with International Solidarity of Scientists) used these specifications to design the first version of SEDNTERP using Visual Basic in the early 1990's.

In the mid 1990's major new features were added to this Visual Basic program by Dr. John S. Philo (jphilo@mailway.com), then at Amgen, and later at Alliance Protein Laboratories.

After useful input from various beta testers, including Geoff Howlett, SEDNTERP 1.00 was then released as a freeware 16-bit Visual Basic 3 program in May of 1997. Subsequently John Philo issued 9 updates through version 1.09 in September 2006.

Then it was decided that Dr. Tom Laue (Tom.Laue@unh.edu) and the BITC Center at the University of New Hampshire would take over development of SEDNTERP and create a new "Version 2.0" that would be written such that (1) it could be run on multiple platforms; (2) it would use public-domain SQL database servers, and (3) other parties could develop and add new modules and capabilities. A "beta" Version 2 of SEDNTERP was first released in 2011, and is still available for download at http://www.jphilo.mailway.com/download . Unfortunately money for developing this new version ran out before it was fully completed, and development of version 2 halted in 2013 at the "beta release" stage, with multiple known bugs still present.

In late 2019 John Philo decided to re-write SEDNTERP from scratch and add new features to make it more useful to both the AUC and light scattering communities. This new version 3 has a user interface and database structure similar to that in version 2.0. It is written as open-source code using the free 'Community' version of Microsoft Visual Studio so hopefully it can be maintained and/or extended by others.

Primary Reference

The features of SEDNTERP 3, the theory and equations behind the properties it computes, and the tables of property values for amino acids, conjugates, and buffer components that it uses in those computations, are all documented in the "primary reference", Reference 97:

> Philo, J. S. (2023). SEDNTERP: a calculation and database utility to aid interpretation of analytical ultracentrifugation and light scattering data. *Eur. Biophys. J.* 52, 233-266.

A previous version of a sedimentation interpretation program is completely described in a Chapter of "Analytical Ultracentrifugation in Biochemistry and Polymer Science" (Ref. 59). That chapter does **not** describe all of the ideas used in, or features currently found in SEDNTERP, but it does give most of the theoretical background and many of the references useful to understand SEDNTERP. In a few cases the formulae used in SEDNTERP differ from those in Ref. 59 (mostly due to correction of typographical errors), and in such cases those differences are noted herein.

Who to Contact for Information and Feedback

For general questions regarding interpretation of sedimentation data and the theory behind this program, contact Dr. Thomas Laue, Professor Emeritus, Department of Biochemistry, University of New Hampshire, Durham, NH 03824. (Tom.Laue@unh.edu)

For problems with program installation or program bugs, contact Dr. John Philo (jphilo@mailway.com)

2.2 What's new in version 3?

- A new user interface, drawing on some elements of the version 2 interface, but with better separation of sample and experimental properties (and full separation of sedimentation equilibrium and sedimentation velocity experiments)
- Support for diffusion experiments (and diffusion coefficients determined from sedimentation velocity)
- Calculation of protein refractive increments and the refractive index of your buffers (features which will be useful to the light scattering community as well as in AUC)
- Concentrations of buffer components can now optionally be entered in molal, mg/mL, g/mL, or weight % units (or volume % for liquid solutes). Different components also can use different concentration units.
- Values for sedimentation (or diffusion) coefficients at multiple concentrations can be entered, fitted to evaluate k_s and $s^0_{20,w}$ (or k_D and $D^0_{20,w}$), and stored in the database
- Experimental molar masses can now be calculated by combining sedimentation and diffusion experiments (which can be in different buffers)
- Hydrodynamic results may now be based on either sedimentation or diffusion data, and may now combine data from experiments run in different buffers (*e.g.* equilibrium and velocity done in different buffers)
- Partial specific volumes calculated from amino acid composition may now optionally include corrections for the volume of the N- and C-terminal residues and the Traube covolume term (refs. 94, 95), which should give more accurate results, particular for peptides.
- Disk (short cylinder) shape models are now supported, and cylinder models are now accurate for length/diameter ratios from 0.01 to 100 (*vs.* only 2 to 20 for versions 1 and 2)
- Hydration can now be calculated for rod and disk models when the asymmetry (*L/d* or *d/H* ratio) is known
- Viscosity can now be calculated for buffers containing D_2O, H_2O^{18} and D_2O^{18}
- The displayed numeric values are now all color-coded to indicate whether the value is a user input (black), calculated (blue), or a user "override" input of a value that is normally calculated (green)
- The graphs of hydrodynamic shape models and of sample properties (titration curves and extinction spectra) are now generated automatically and shown within the main program form.
- 53 tables of refractive index *vs.* concentration for existing solutes were fitted to polynomials and the resulting polynomial coefficients added to the database. Six new solutes of particular interest to the biotechnology community were also added (L-arginine, L-alanine, L-histidine, L-glycine, sorbitol, and Tween-80), all of which have both density and viscosity data (and all but sorbitol and L-glycine have refractive index data also).

- The default *pKa* values for the amino acids and the N- and C-terminus have been updated to reflect newer experimental values from Thurkill *et al.* (2006), Ref. 93.
- The equations for the density and viscosity of water have been updated to reflect current "recommended" values and changes to the International Practical Temperature Scale. The values of the fundamental constants have also been updated to current recommended values. (Yes these things do change over the ~25 years since SEDNTERP was created!)
- This document has been completely revised and re-formatted, and 28 additional references added.

What's missing?

- "Loosely-bound" conjugates ("non-covalent" conjugates) are no longer supported. All conjugates are assumed to be fully bound (either covalently or non-covalently) and therefore are included when computing the total molar mass, overall partial specific volume, overall refractive increment, *etc.*
- Pasting or entering amino acid sequences in 3-letter AA codes is no longer supported.
- Saving or reading sample composition or buffer composition files (.smp and .buf files) is no longer supported. (Those features were a hold-over from pre-release versions that did not use a database.)

2.3 List of Parameters for Input

The table below summarizes the sample, solution, and molecular parameters that can be input or calculated, with links to the theory or background section(s) for that parameter, and links to the program form(s) or dialogue box(es) where that parameter is entered or displayed. The parameters are listed roughly in the order they appear on the 4 major tabs of the Main form.

Parameter	Theory or Background	Tab or Form
M or M_r ; molar mass (can be entered or computed)	Calculating Molar Mass	Sample Properties major tab Sample Composition form Sedimentation Equilibrium minor tab Sedimentation Velocity

		minor tab
$\bar{v}$; partial specific volume	Partial Specific Volume Buoyant Molar Mass Calculating Vbar Estimates	Sample Properties major tab Sample Composition form Sedimentation Equilibrium minor tab Sedimentation Velocity minor tab
δ_1 ; hydration	Hydration and Asymmetry Limitations on Estimating Hydration and Asymmetry	Sample Properties major tab Sample Composition form Derived Hydrodynamic Results major tab
k_s ; concentration dependence of the sedimentation coefficient	Concentration Dependence of the Sedimentation Coefficient	Sample Properties major tab Sedimentation Velocity minor tab Evaluate ks and Extrapolate s0(20,w) form
k_D ; concentration dependence of the diffusion coefficient	Concentration Dependence of the Diffusion Coefficient	Sample Properties major tab Diffusion minor tab Evaluate kD and Extrapolate D0(20,w) form

dn/dc ; specific refractive index increment	Calculating Refractive Increment Estimates Interpolating a Buffer Refractive Index	Sample Composition form Buffer Composition form
ρ ; buffer density	Interpolating a Buffer Density Buoyant Molar Mass Converting s* to s0(20,w)	Buffer Properties major tab Buffer Composition form Sedimentation Velocity minor tab Sedimentation Equilibrium minor tab
η ; buffer viscosity	Interpolating a Buffer Viscosity Converting s* to s0(20,w) Converting D* to D0(20,w)	Buffer Properties major tab Buffer Composition form Sedimentation Velocity minor tab Diffusion minor tab
n ; buffer refractive index	Interpolating a Buffer Refractive Index Calculating Refractive Increment Estimates	Buffer Properties major tab Buffer Composition form
*s** ; apparent or 'raw' sedimentation coefficient	s*, Apparent Sedimentation Coefficient Converting s* to s0(20,w)	Sedimentation Velocity minor tab
*D** ; apparent or 'raw' diffusion	D*, Apparent Diffusion	Diffusion minor tab

coefficient	Coefficient Converting D* to D0(20,w)	Sedimentation Velocity minor tab
σ ; sigma, reduced apparent molar mass	What is Sigma (and Why is it Useful)? Calculating Molar Mass	Sedimentation Equilibrium minor tab
RPM ; revolutions per minute	Calculating Molar Mass	Sedimentation Equilibrium minor tab
Temperature	Calculating Vbar Estimates Calculating Refractive Increment Estimates Interpolating a Buffer Density Interpolating a Buffer Viscosity Interpolating a Buffer Refractive Index Calculating Molar Mass Converting s* to s0(20,w) Converting D* to D0(20,w)	Buffer Properties major tab Sedimentation Velocity minor tab Sedimentation Equilibrium minor tab Diffusion minor tab Sample Composition form
c ; sample weight concentration (mg/mL)	Converting s* to s0(20,w) Converting D* to D0(20,w)	Sedimentation Velocity minor tab Diffusion minor tab Evaluate ks and Extrapolate s0(20,w) form Evaluate kD and Extrapolate D0(20,w)

		form
a/b, L/d or *d/h;* ellipsoidal or cylindrical asymmetry	Hydration and Asymmetry Limitations on Estimating Hydration and Asymmetry Calculating Ellipsoidal Asymmetry Calculating Cylindrical Asymmetry	Derived Hydrodynamic Results major tab

2.4 List of Interpolations Available

The table below lists properties that SEDNTERP will interpolate. For each property, it lists which user inputs are needed to make that interpolation, a link to the corresponding page where the details of that calculation are described, and a link to the form where that property is calculated or displayed.

Property	**Input**	**Calculation Details**	**Form**
Vbar	amino acid composition, conjugates, temperature	Calculating Vbar Estimates	Sample Composition form
Density	buffer components and concentrations; % isotopes of water; temperature	Interpolating a Buffer Density	Buffer Composition form
Viscosity	buffer components and concentrations; % isotopes of water; temperature	Interpolating a Buffer Viscosity	Buffer Composition form
Refractive Index	buffer components and concentrations; % isotopes of	Interpolating a Buffer Refractive	Buffer Composition

Property	Input	Calculation Details	Form
	water; temperature; wavelength	Index	form
Hydration δ_1	amino acid composition; pH	Estimating Hydration	Sample Composition form
Isoelectric Point	sample composition	Calculating Electric Charge and Isoelectric Point	Sample Composition form
Electric Charge	sample composition	Calculating Electric Charge and Isoelectric Point	Sample Composition form
Absorbance, extinction coefficient	sample composition	Calculating UV Absorbance Properties	Sample Composition form
Specific refractive index increment	sample composition; buffer refractive index; wavelength	Calculating Refractive Increment Estimates	Sample Composition form

2.5 List of Results Available

"Results" are molecular parameters computed by SEDNTERP from the equations described in the theory section. To understand the results, the theory section ought to be read at length. The table below shows which parameters affect each result and gives links to the corresponding theory or background topic(s) and links to the form (or tab on the main form) on which that result is

computed or displayed.

Result	Dependence	Theory or Background	Form or Tab
M_r molar mass	If from sedimentation equilibrium: Depends on sigma, RPM, sample vbar, buffer density, and temperature If from composition: Depends on user entered composition, and database of amino acid and conjugate masses	What is Sigma (and Why is it Useful)? Calculating Molar Mass	Sample Properties major tab Sample Composition form
M_b buoyant molar mass	Depends on molar mass, partial specific volume, and buffer density	Buoyant Molar Mass Calculating Buoyant Molar Mass	Sedimentation Equilibrium minor tab
$s_{20,w}$ Sedimentation coefficient corrected to 20 degrees C in water	Depends on apparent *s (s*)*, temperature, buffer density, buffer viscosity, and vbar	Converting s* to s0(20,w)	Sedimentation Velocity minor tab Evaluate ks and Extrapolate s0(20,w) form
$s^0_{20,w}$	Depends on apparent *s (s*)*, temperature,	Concentration Dependence of	Main Data Entry formSedimentation

Sedimentation coefficient corrected to 20 degrees C in water and extrapolated to zero solute concentration	sample concentration, k_s, buffer density, buffer viscosity, and vbar	the Sedimentation Coefficient Converting s* to s0(20,w)	Velocity minor tab Evaluate ks and Extrapolate s0(20,w) form
$D_{20,w}$ Diffusion coefficient corrected to 20 degrees C in water	Depends on apparent *D (D*)*, temperature, buffer density, buffer viscosity, and vbar	Converting D* to D0(20,w)	Diffusion minor tab
$D^0_{20,w}$ Diffusion coefficient corrected to 20 degrees C in water and extrapolated to zero solute concentration	Depends on apparent *D (D*)*, temperature, sample concentration, k_D, buffer density, buffer viscosity, and vbar	Concentration Dependence of the Diffusion Coefficient Converting D* to D0(20,w)	Diffusion minor tab Evaluate kD and Extrapolate D0(20,w) form
R_0 or R_p Equivalent radius	Depends only on molar mass if Teller's empirical formula is used (R_p); Depends on molar mass and vbar otherwise (R_0)	Calculating an Equivalent Radius	Derived Hydrodynamic Results major tab

f experimental frictional coefficient	Depends on $s^0_{20,w}$ (if computed from sedimentation) or on $D^0_{20,w}$ (if computed from diffusion)	Calculating the Standardized Experimental Frictional Coefficient, f	Derived Hydrodynamic Results major tab
R_s Stokes radius	Depends on $s^0_{20,w}$ (if computed from sedimentation) or on $D^0_{20,w}$ (if computed from diffusion)	Calculating the Stokes Radius	Derived Hydrodynamic Results major tab
f_0 or f_p minimum frictional coefficient	Depends on equivalent radius	Calculating the Minimum Frictional Coefficient	Derived Hydrodynamic Results major tab
S_{max} maximum sedimentation coefficient	Depends on minimum frictional coefficient, molar mass, vbar, and temperature	Calculating the Minimum Frictional Coefficient	Derived Hydrodynamic Results major tab
D_{max} maximum diffusion coefficient	Depends on minimum frictional coefficient and temperature	Calculating the Minimum Frictional Coefficient	Derived Hydrodynamic Results major tab
$\delta_{1,max}$ maximum hydration of a molecule	Depends on minimum frictional coefficient, and $s^0_{20,w}$ or $D^0_{20,w}$	Calculating Maximum Hydration	Derived Hydrodynamic Results major tab

f_{hyd}/f_0 or f_{hyd}/f_p frictional ratio of hydrated vs. anhydrous equivalent sphere	Depends on hydration (δ_1) and equivalent radius	Estimating Hydration	Derived Hydrodynamic Results major tab
P Perrin translational friction function	Depends on frictional coefficient (*f*), hydration (δ_1) and equivalent radius	Calculating the Standardized Experimental Frictional Coefficient, f	Derived Hydrodynamic Results major tab
maximum asymmetry of a molecule $(a/b)_{max}$ or $(L/d)_{max}$	Depends on shape model (ellipsoid or cylinder), $s^0_{20,w}$ or $D^0_{20,w}$, and minimum frictional coefficient	Calculating Ellipsoidal Asymmetry Calculating Cylindrical Asymmetry	Derived Hydrodynamic Results major tab
estimated shape and length of a molecule	Depends on shape model (ellipsoid or cylinder), $s^0_{20,w}$ or $D^0_{20,w}$, minimum frictional coefficient, and hydration	Calculating Ellipsoidal Asymmetry Calculating Cylindrical Asymmetry	Derived Hydrodynamic Results major tab
δ_1 hydration, calculated from ellipsoidal or cylindrical asymmetry model	Depends on $s^0_{20,w}$ or $D^0_{20,w}$, minimum frictional coefficient, assumed asymmetry ratio, and shape model (prolate,	Calculating Ellipsoidal Asymmetry Calculating Cylindrical Asymmetry	Derived Hydrodynamic Results major tab

	oblate, rod, disk)		

2.6 List of Symbols Used

a/b	axial ratio (ratio of major to minor axis) for an ellipsoid model of a molecule
$(a/b)_{max}$	maximum possible axial ratio (the value obtained assuming the molecule has zero hydration)
dn/dc	specific refractive index increment; change in solution refractive index per change in weight concentration
D	diffusion coefficient
$D_{20,w}$ or *D(20,w)*	diffusion coefficient corrected to standard conditions (20 °C, water), but not extrapolated to zero concentration
$D^0_{20,w}$ or *D*0(20,w)	diffusion coefficient corrected to standard conditions (20 °C, water), and also extrapolated to zero concentration
D_{max}	maximum possible diffusion coefficient (the value for an anhydrous sphere with the same molecular volume, diffusing in water at the experimental temperature)
δ_1	hydration, the grams of bound water per gram of protein
$\delta_{1,max}$	maximum possible hydration, calculated from the observed f/f_0 or f/f_p ratio
f	the hydrodynamic frictional coefficient of a molecule

f_0 or f_p	the minimum frictional coefficient for a molecule (calculated assuming it is a perfect sphere and is unhydrated), from the 'vbar' or 'Teller' methods, respectively
f/f_0 or f/f_p	The ratio of the actual molecular frictional coefficient to that an anhydrous sphere, as calculated from the 'vbar' or 'Teller' methods, respectively. This ratio is commonly cited as a measure of hydrodynamic compactness or molecular asymmetry.
f_{hyd}/f_0 or f_{hyd}/f_p	The ratio of the frictional coefficients for a hydrated equivalent sphere (the anhydrous sphere plus a layer of water consistent with the δ_1 hydration parameter) to that for an anhydrous sphere (from either the Teller or vbar methods). This ratio separates the hydration dependence within the f/f_0 ratio from the molecular shape dependence.
ϕ	volume fraction of solute; $\phi = \nu_s c$
ϕ_2	apparent isopotential partial specific volume
η	viscosity (usually the buffer viscosity)
k_s	coefficient for the concentration dependence of the sedimentation coefficient (sometimes called a "non-ideality coefficient")
k_D	coefficient for the concentration dependence of the diffusion coefficient (sometimes called the "diffusion interaction parametner")
L/d	length to diameter ratio for a cylinder model of a molecule
$(L/d)_{max}$	maximum possible length/diameter ratio (the value obtained assuming the molecule has zero hydration)

M_b or *Mb*	buoyant molar mass
M, M_r or *Mr*	molar mass (also known as 'relative' molar mass)
n	buffer refractive index
N_0	Avogadro's number
P	Perrin translational friction function; the ratio of the actual frictional coefficient *f* to that for a hydrated sphere of equal volume (f_{hyd}). This parameter separates the molecular shape dependence within the f/f_0 ratio from the hydration dependence.
ρ	density (usually the buffer density)
ρ_{hyd}	density of the bound water of hydration
R_0 or R_p	equivalent radius (from the 'vbar' or 'Teller' methods, respectively)
R_s	Stokes radius
σ	reduced apparent molar mass
s	sedimentation coefficient
*s**	experimental sedimentation coefficient (not corrected to standard conditions)
$s_{20,w}$ or *s*(20,w)	sedimentation coefficient corrected to standard conditions (20 °C, water), but not extrapolated to zero concentration
$s^0_{20,w}$ or *s*0(20,w)	sedimentation coefficient corrected to standard conditions (20 °C, water), and also extrapolated to zero concentration

S_{max}	maximum possible sedimentation coefficient (the value for an anhydrous sphere with the same molecular volume, diffusing in water at the experimental temperature)
$\bar{\nu}$ or "vbar"	partial specific volume; increase in solution volume per gram of added solute
$\bar{\nu}_p$	Teller "packing" value for $\bar{\nu}$ from crystal structures = 0.7654 mL/g (Ref. 33)
ν_s	"swollen" specific volume (including bound or entrained solvent)
ω	angular velocity [$\omega = 2\pi \times (RPM/60)$]

2.7 How to Cite this Program

The features of SEDNTERP 3, the theory and equations behind the properties it computes, and the tables of property values for amino acids, conjugates, and buffer components that it uses in those computations, are all documented in the "primary reference", Reference 97:

> Philo, J. S. (2023). SEDNTERP: a calculation and database utility to aid interpretation of analytical ultracentrifugation and light scattering data. *Eur. Biophys. J.* 52, 233-266.

The recommended citing procedure is to state in the *Methods* section of your manuscript which calculations were done using "SEDNTERP version 3.xx", and then cite the above reference (Reference 97),

Note that a previous version of a sedimentation interpretation program is completely described in a Chapter of "Analytical Ultracentrifugation in Biochemistry and Polymer Science" (Ref. 59). That chapter does **not** describe all of the ideas used in, or features currently found in, version 3 of SEDNTERP, nor does it include as authors either John Philo or David Hayes, who created the code.

2.8 Acknowledgements

Work on the original project was supported in part by a Grant-in-Aid #850-781 and #870-1092 from the American Heart Association and Grants BBS 86-15815 and DIR 90-02027 from the National Science Foundation to Thomas Laue.

Subsequent work was continued under NSF grant #9314040 with the help of an ROA grant attached to the original grant.

John Philo is thanked for developing major new features for this program, and he and Amgen are thanked for allowing them to be included in the original distribution of the program. John is also thanked for continuing to revise and update the program and this Help file.

David Hayes is thanked for his continued interest and support of the program and for recently finding a very subtle bug.

Mattia Rocco is thanked for pushing for the addition of the buffer refractive index and *dn/dc* calculations, the addition of the terminal group and covolume terms to the protein $\bar{\nu}$ calculations, and his help with debugging.

The references listed in this Help file acknowledge the irreplaceable heritage of contributions to science from people around the world.

Finally, the three authors wish to thank the many users who have told us SEDNTERP is useful and have encouraged us to keep this project going.

2.9 Glossary

A

asymmetry
Asymmetry refers to whether a molecule is compact and globular in shape versus highly extended (such as a rod-like molecule or a random coil).

B

buffer
In this document "buffer" means the solution in which the protein or other macromolecule is dissolved (whether or not that solution contains a buffering agent). That is, the buffer represents all solution components other than the macromolecule.

C

conjugate
A co-factor, protein glycan, or other molecule bound to a protein, either covalently or non-covalently, such that it should be included when calculating the total mass, partial specific volume, *etc*.

E

equilibrium
The state in which opposing forces or processes are in a balance such that no overall change occurs in the system. In sedimentation equilibrium experiments the flow of a solute due to the gravitational field produced by the ultracentrifuge is balanced by the flow due to diffusion. This balance is measured by sigma (σ), the exponent of the slope of a stable concentration gradient.

equivalent radius
The equivalent radius is the radius of an anhydrous sphere with a volume equivalent to that of a given protein or other macromolecule.

H

heterogeneity
Heterogeneity refers to the presence of more than one species of molecule. This term is used most often when it could be assumed or normally happens that one species is present. For example, if an enzyme protein solution contains an unmodified enzyme species and an enzyme species with a post-translational carbohydrate chain added to it, the solution will display heterogeneity in its apparent molar mass even though it is not necessarily heterogeneous in its enzymatic activity.

hydration
Hydration refers to water molecules which sediment or diffuse along with a macromolecule, either because they are bound to its surface, or because they are physically entrained due to the molecule's shape.

I

interpolate
In the program SEDNTERP, **interpolate** means to estimate physical parameters from standard tables and user-entered information. In most cases, SEDNTERP uses a power series approximation of standard tabular data. For example, SEDNTERP will interpolate the density of a solution from the concentration of each solute entered by the user.

interpolation

In the program SEDNTERP, **interpolation** refers to the estimation of physical parameters from standard tables and user-entered information. In most cases, SEDNTERP uses a power series approximation of standard tabular data. For example, SEDNTERP will interpolate the density of a solution from the concentration of each solute entered by the user.

interpretation
In the program SEDNTERP, **interpretation** refers to the computation of molecular structural properties from experimental values (such as sedimentation coefficients and/or molar masss) and from experimental conditions (such as buffer density).

P

Perrin P
The ratio of the translational frictional coefficient for an anhydrous molecule to that for an anhydrous sphere with the same volume. In other words, *P* is the f/f_0 ratio for the anhydrous molecule.

S

sample
In the program SEDNTERP, the word **sample** is used to denote the molecular species that is the subject of an sedimentation experiment. Only the macromolecule being sedimented is called the sample.

sedimentation
Sedimentation is the general term for the movement of solutes caused by a centrifugal field in the ultracentrifuge.

sedimentation coefficient
The **sedimentation coefficient**, **s**, has units of seconds, and is often expressed in units of Svedbergs (S), where 1 S = 10^{-13} sec. It is a measure of how fast a molecule sediments in a centrifugal field. It is determined experimentally in a sedimentation velocity experiment.

Stokes radius
The Stokes radius of a molecule is defined as the radius of an unhydrated sphere that would have the same diffusion coefficient as that molecule.

3 Theory of Computer Aided Interpretation of Sedimentation Data

3.1 Background

3.1.1 Introduction

The theory described here focuses on the interpretation of sedimentation and/or diffusion data for purified proteins, glycoproteins or lipoproteins exhibiting relatively simple behavior in aqueous buffers. Proper interpretation of AUC data can provide estimates of molar mass, subunit stoichiometries, association constants, thermodynamic non-ideality, molecular sizes, and relative structural asymmetry for a wide variety of macromolecules. Estimates of the reduced apparent molar mass (σ) or the apparent sedimentation coefficient (s^*) only require knowledge of the concentration distribution, rotor speed, and the time from the start of the experiment (Refs. 3, 8, 9, 10, 11, 12). However, full interpretation of sedimentation or diffusion data requires knowledge of ancillary information such as the solution density and viscosity, as well as the protein's partial specific volume, $\bar{\nu}$. Such data may be entered into SEDNTERP directly (and temperature corrections computed), or the program may use various methods to interpolate estimates of this essential ancillary information from simpler information when explicit data are not available. In this document, the equations used for interpretation and interpolation are presented (with references), and the limitations of these computations are described.

Naturally, proper data interpretation requires good data acquisition. Many practical suggestions are available in books (Refs. 8, 10), operating manuals (Refs. 11, 12) or reviews (Ref. 9) that describe the acquisition of sedimentation velocity (Refs. 9, 10, 11), sedimentation equilibrium (Refs. 9, 10, 11, 12, 13, 14) and diffusion data (Refs. 9, 10). The approaches implemented by SEDNTERP are appropriate starting points for interpreting data from these experiments, but the interpretations computed by SEDNTERP are not comprehensive enough to deal with the sedimentation of non-proteins or complex chemical systems (Refs. 15, 16, 17).

3.1.2 Equations for s*

The apparent sedimentation coefficient *s is determined directly from the experimental data:**

The apparent sedimentation coefficient s^* is defined as the velocity of a sedimenting particle

divided by the acceleration. Simple physics (mostly Newton's first law) tells us that the sedimentation coefficient can be determined from the slope of a graph of *ln(r)* versus $\omega^2 t$, where *r* is the distance from the center of rotation to the midpoint of the sedimenting concentration boundary, ω is the angular velocity of the rotor, and *t* is time. That graphical approach is still valid, but today we instead typically determine *s** via computerized analysis of many velocity scans, using methods such as time-derivative (*dc/dt*) analysis (ref. 60), or direct fitting of the scans to solutions of the Lamm equation (Refs. 78, 79).

The interpretation of sedimentation coefficients is based upon

Equation 4:

$$s^* = \frac{M_b}{N_0 f} = \frac{M\left(1-\bar{v}\rho\right)}{N_0 f}$$

where *M* is the solute's anhydrous molar mass (g/mole), $\bar{v}$ is its partial specific volume (mL/g), ρ is the solvent density (g/mL), N_0 is Avogadro's number, M_b is the buoyant molar mass and *f* is the translational frictional coefficient. Note that Equation 4 can be derived by combining Eq. 3 for the diffusion coefficient with the Svedberg equation (Eq. 2).

For a hydrodynamically ideal molecule, all information concerning the shape and size of the molecule is expressed through *f*. Changes in *f* are propagated inversely to *s*. Equation 3 shows that knowledge of *M* and *s* are sufficient to determine *f*, even in the absence of determining *D*. This latter approach to the determination of *f* is often preferred because accurate values of *M* and *s* are generally easily obtained, but accurate determinations of *D* can be difficult.

Note: For solutions containing two or more non-interacting proteins, and in cases where the boundaries are well resolved, the analysis of *s** described above, and the other calculations described in this manual such as analysis of asymmetry, may in principle be applied to each component's boundary. It must be kept in mind, however, that the effects of solute-solute interaction are not always innocuous and they can produce measurable alterations in the shape and position of the boundaries (Ref. 20).

3.1.3 Equations for D*

The diffusion coefficient *D** can be related to several other molecular parameters. It is directly related to the translational frictional coefficient, *f*, through the Einstein-Smoluchowski relationship (Ref. 107):

Equation 3:

$$D^* = \frac{RT}{N_o f} = \frac{k_B T}{f}$$

where *R* is the gas constant, *T* is temperature (Kelvin), N_o is Avogadro's number, and k_B is the Boltzmann constant.

The right-hand term of Equation 3 was incorrect in the published version of the primary reference (ref. 97) due to inadvertent deletion of the "*T*".

It is worth noting that this relation means that *f* (and therefore information about hydration and asymmetry) can be obtained from *D** without needing to know either the partial specific volume or the molar mass (both advantages relative to the sedimentation coefficient).

The diffusion coefficient is also directly related to the Stokes radius (hydrodynamic radius), R_s, through the relation:

$$R_s = \frac{k_B T}{6\pi\eta D^*}$$

where η is the buffer viscosity.

3.1.4 What is Sigma (and Why is it Useful)?

In sedimentation equilibrium experiments, the reduced apparent molar mass, σ, is the quantity which is directly measured by experiment. Sigma measures the exponential curvature of the equilibrium concentration gradient.

By equating the mass flow due to sedimentation to that arising from diffusion, one obtains the fundamental equation describing sedimentation equilibrium of a single, thermodynamically ideal protein:

Equation 14:

$$\sigma = \frac{\partial \ln(c(r))}{\partial\left(r^2/2\right)} = \frac{s\omega^2}{D} = \frac{M_b\omega^2}{RT} = \frac{M\left(1-\bar{v}\rho\right)\omega^2}{RT}$$

where σ is the reduced apparent molar mass (Refs. 14, 25), *c(r)* is the weight concentration at radial position *r*, and the other terms are as described in the table of symbols.

While σ may be determined from the slope of a graph of ln(*c*) as a function of $r^2/2$, sedimentation equilibrium data is best analyzed by nonlinear least squares procedures, which can accommodate more complicated cases involving association and non-ideality (Refs. 2, 10, 14, 26).

To find the molar mass from sigma, SEDNTERP uses the rightmost expression in Equation 4. That means it is necessary to know the conditions of the experiment (temperature and rotor speed), and for accurate molar masses the density of the solution and the partial specific volume of the protein are also needed. SEDNTERP will also calculate and display sigma from an entered buoyant or true molar mass provided a rotor speed has been entered.

Why is sigma useful?

Sigma units for the molar mass are included on the Sedimentation Equilibrium minor tab partly for compatibility with the sedimentation equilibrium data analysis program NONLIN (which only works in sigma units), but also because they are quite useful in interpreting and designing SE experiments. For example, a sigma above 5 means the concentration gradient is likely too steep to measure accurately by absorbance optics, while a sigma near 1 means the gradient is very shallow, and the returned molar mass will depend very strongly on knowing the zero offset accurately.

3.1.5 Buffer and Sample Properties Needed for Experiment Interpretation

The buffer density & viscosity, and the sample partial specific volume, are usually needed for further interpretation of s*.

*s** can be directly derived from experimental data. However, more often than not, it is the frictional coefficient *f* and/or the molar mass *M* that are of interest to the experimenter, since these are required to learn about the size and shape of a molecule.

Equation 3:

$$s = \frac{M(1-\bar{v}\rho)}{N_0 f} = \frac{M_b}{N_0 f}$$

Equation 3 above shows that $\bar{v}$ and ρ are required to determine *f* from *M* and *s*.

Only the buffer viscosity is needed for interpreting D*

Equation 7 below shows that *f* can be determined from *D** without knowledge of $\bar{v}$ or ρ.

Equation 7:

$$D^* = \frac{RT}{N_o f} = \frac{k_B T}{f}$$

However, the experimental *f* that appears in equations 3 and 7 is not an intrinsic molecular property, but rather one that implicitly depends on the buffer viscosity (it is proportional to the viscosity). Therefore before interpreting *f* to estimate molecular hydration and/or asymmetry it is necessary to correct *f* to the value it would have under standard conditions (in water), which in turn means we need to know the buffer viscosity.

Further interpretation of *f* requires that estimates of *f* be made for anhydrous spheres of identical molar mass and density. These calculations permit, through still more calculations, the estimation of the limits on a macromolecule's hydration and asymmetry. While none of the calculations is difficult, SEDNTERP assembles all of them into one program to facilitate data interpretation. This not only eliminates the tedium associated with routine analysis, it also provides a consistent and reliable means of evaluating hydrodynamic observations, and storing your experiment information in the program database allows you to quickly reproduce that evaluation (and archive results used in reports or publications).

If the molar mass was obtained from sedimentation equilibrium then that also requires knowing vbar and the buffer density

Today we typically know the molar mass of a protein either from its amino acid composition or from mass spectroscopy. However if and when it is necessary to determine *M* via sedimentation equilibrium,

Equation 6:

$$M_b = M\left(1-\bar{v}\rho\right)$$

Equation 6 above shows that calculation of M_r from M_b requires knowing both $\bar{v}$ and ρ.

3.1.6 Concentration Dependence of the Sedimentation Coefficient

In the absence of reversible-self association, sedimentation rates normally decrease as the concentration of the macromolecule increases, due to multiple complex "hydrodynamic non-ideality" effects. A major cause of this non-ideality is that solvent must flow in the direction opposite to sedimentation to fill the space formerly occupied by the sedimenting macromolecule, and thus the molecules must swim 'upstream' against this solvent backflow. The velocity of that backflow is proportional to the volume fraction occupied by the macromolecule, and thus increases with macromolecule concentration.

The extent to which the sedimentation rate is reduced at a given concentration depends on the molecular conformation. In general the reduction will be greater for molecules that are highly asymmetric or flexible.

The concentration dependence of the sedimentation coefficient is usually characterized by an empirical parameter $\boldsymbol{k_s}$, which is essentially defined by the relation

Equation 5:

$$s(c) = \frac{s_0}{1 + k_s \cdot c}$$

where $s(c)$ is the sedimentation coefficient (either raw or standardized) at finite weight concentration c, and s_0 the value that would be observed at infinite dilution. Note that the above relation typically does describe experimental results up to moderate concentrations (perhaps 10-20 mg/mL), but it is essentially a first-order approximation that may not hold at concentrations where the sedimentation coefficient is strongly reduced.

The parameter k_s has units of inverse concentration, and thus when c is in units of mg/mL it can have units of mL/mg. However, in those units its numerical value is quite small (~0.01 for compact, globular proteins). Therefore it is often more convenient to use units of mL/g for k_s , which usually makes the numerical value roughly >= 5. Using units of mL/g also facilitates comparison to theoretical values based on multiples of the specific volume (as will be discussed below).

How is k_s measured?

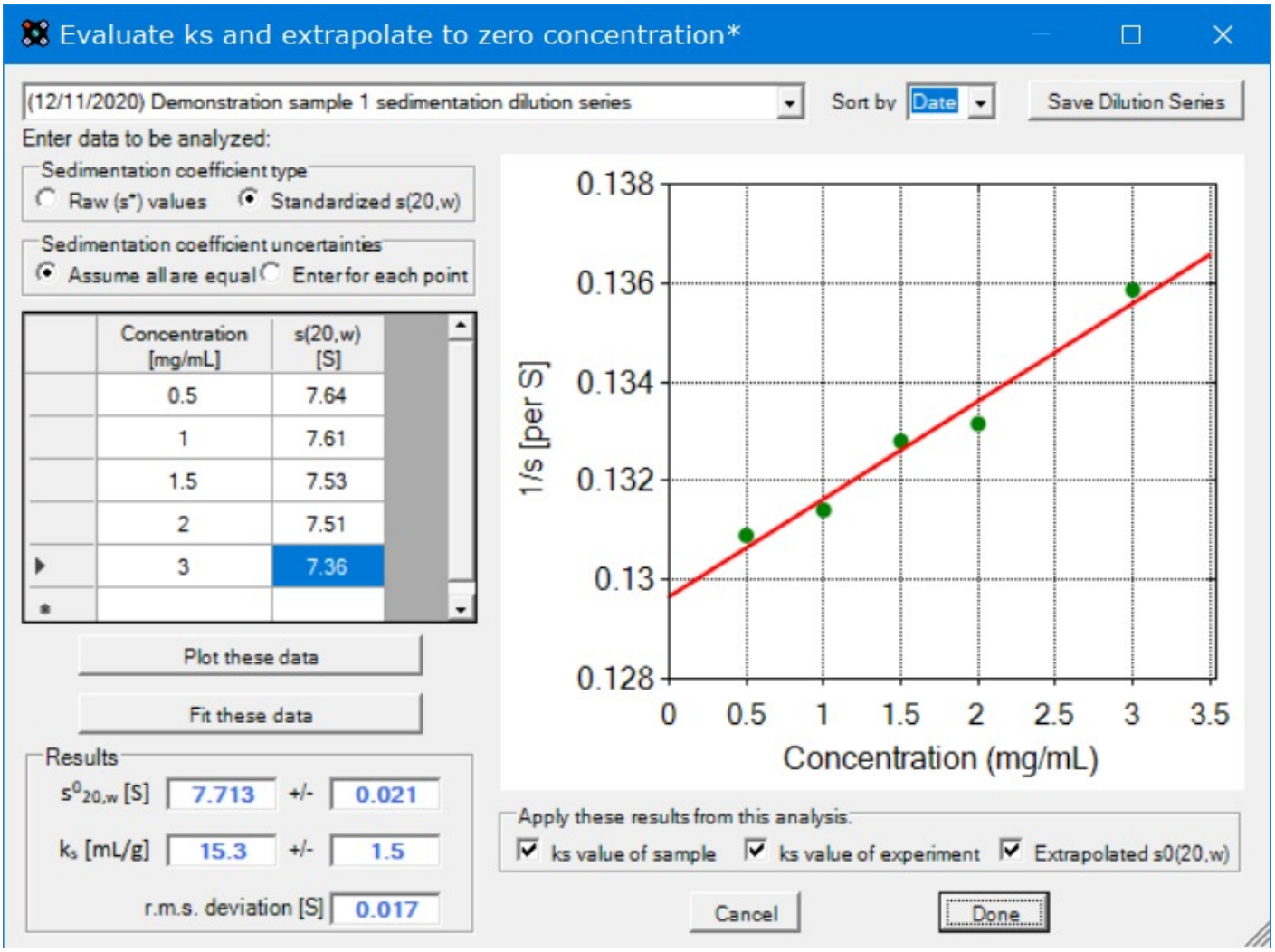
Determining an experimental value for k_s requires measuring the sedimentation coefficient at

several concentrations, and then fitting those results to Eq. 5 above. Re-arranging Eq. 5 we get this expression

Equation 12:

$$\frac{1}{s(c)} = \frac{1}{s_0}(1 + k_s \cdot c)$$

Equation 12 shows that if we plot $1/s(c)$ versus concentration, we should get a straight line whose slope is k_s / s_0. Furthermore, the intercept of that line gives us the inverse of the sedimentation coefficient at infinite dilution. Fitting a set of experimental values to the above equation is implemented in SEDNTERP via the Evaluate ks and Extrapolate s0(20,w) form, shown below:



Can the value for k_s be predicted?

Unfortunately the magnitude of the hydrodynamic non-ideality cannot be predicted based on thermodynamic principles. However hydrodynamic theory and numerical methods have shown that k_s is directly related to the "swollen" specific volume, ν_s , which includes the bound or entrained solvent.

Various published theoretical or numerical simulation studies indicate that k_s can be calculated as a multiple of ν_s between 4 and 7.16 (see ref. 73 for a review of theoretical studies prior to 1985). Among the theoretical studies that of Batchelor (ref. 71) is often considered to be the best. That study gives $k_s = 6.55\nu_s$. On the other hand, more recent Brownian dynamics simulations by Brady and Durlofsky (ref. 72) yield $k_s = 5\nu_s$.

Those expressions then beg the question "what is the value for the 'swollen' specific volume?". At a minimum, that value should be the sum of the partial specific volume, $\bar{\nu}$, and the volume of the water of hydration. Thus for a typical protein with $\bar{\nu}$ = 0.73 mL/g and δ_1 = 0.4 g/g we expect ν_s ~= 1.13 mL/g and therefore k_s = 5.7 mL/g (ref. 71) to 7.4 mL/g (ref. 72). Those values are however distinctly lower than the value of ~9 mL/g reported for compact globular proteins (refs. 21, 22, 27, 28). (Note that this experimental value of 9 mL/g is actually the default k_s value assumed by SEDNTERP when the program is installed.) That is, it appears the swollen volume that applies to calculating k_s is usually at least somewhat larger than expected based on the volume for the protein plus its "bound" water. Indeed, a review of the literature by Creeth & Knight (1965) indicates that for globular or moderately asymmetric proteins the swollen volume is ~1.6-3.5 times larger than $\bar{\nu}$, while for extremely extended proteins like myosin or collagen ν_s can be > 100 times $\bar{\nu}$ (ref. 90). The swollen volume is often taken as being the 'hydrodynamic' volume implied by the Stokes radius (*i.e.* the volume of a sphere with radius = R_s), which again can imply a swollen specific volume considerably larger than $\bar{\nu}$.

It should be noted that it has recently (2023-24) been shown that the swollen volume, hydration, and other hydrodynamic properties of proteins (Ref. 102), and even of flexible polymers such as PEG (Ref. 104), can be accurately calculated from their atomic models by a program called HullRad.

3.1.7 Concentration Dependence of the Diffusion Coefficient

In the absence of reversible-self association, apparent diffusion rates typically increase as the concentration of the macromolecule increases, although in some cases they will instead decrease

with concentration. These changes are due to both thermodynamic non-ideality and complex hydrodynamic non-ideality effects.

The extent to which the apparent diffusion coefficient increases at a given concentration may depend strongly on the molecules electric charge, the ionic strength, and interactions between the macromolecule and the solvent.

The concentration dependence of the diffusion coefficient is usually characterized by an empirical parameter $\boldsymbol{k_D}$, which is essentially defined by the relation

Equation 6:

$$D(c) = D_0\left(1 + k_D \cdot c\right)$$

where $D(c)$ is the diffusion coefficient (either raw or standardized) at finite weight concentration c, and D_0 the value that would be observed at infinite dilution.

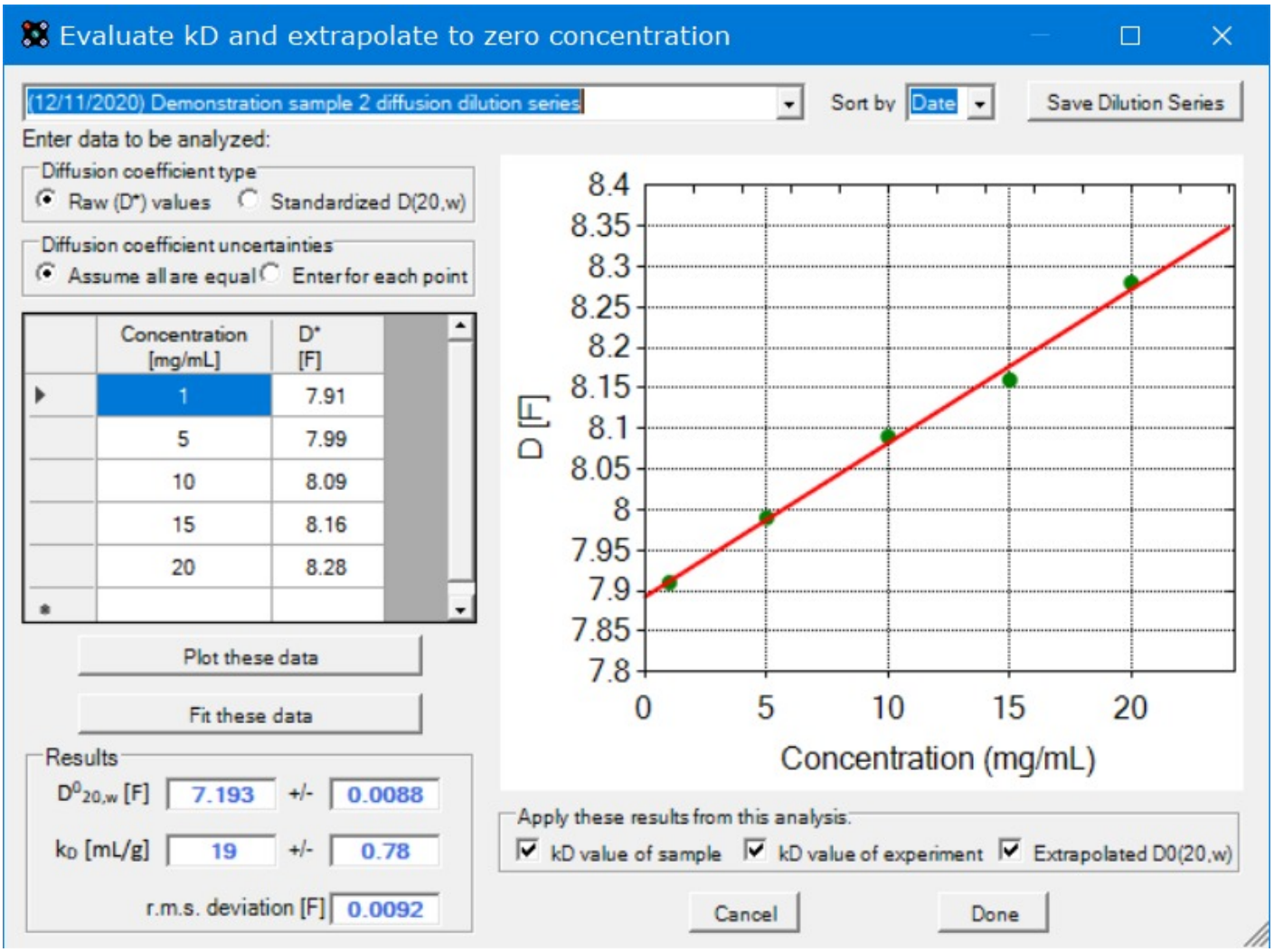
Note that the parameter k_D is also known as the "diffusion interaction parameter", especially when it is measured using dynamic light scattering.

Note also that Equation 6 above typically does describe experimental results up to moderate concentrations (perhaps 10 mg/mL), but it is essentially a first-order approximation that may not hold at concentrations where the diffusion coefficient is substantially increased or decreased.

The parameter k_D has units of inverse concentration, and thus when c is in units of mg/mL it can have units of mL/mg. However, in those units its numerical value may be quite small (~0.004 for globular proteins with relatively low electric charge, ref 74). Therefore it is often more convenient to use units of mL/g for k_D, which usually makes the numerical value in the range of 1-100 (positive or negative). Using units of mL/g also facilitates comparison to theoretical values based on multiples of the specific volume (as will be discussed below).

How is k_D measured?

Determining an experimental value for k_D requires measuring the diffusion coefficient at several concentrations, and then fitting those results to **Eq. 6** above. That equation shows that if we plot $D(c)$ versus concentration, we should get a straight line whose slope is k_D. Furthermore, the intercept of that line gives us the diffusion coefficient at infinite dilution. Such a plot and fitting procedure are implemented in SEDNTERP on the Evaluate kD and Extrapolate D0(20,w) form, shown below:



Note that the value for k_D may also depend on the technique used to measure D. In particular, diffusion methods that generate a concentration gradient (such as AUC experiments in a synthetic boundary cell, a standard sedimentation velocity run, or a Taylor dispersion measurement) are usually carried out under conditions of constant temperature and constant chemical potential, whereas diffusion coefficients obtained using dynamic light scattering are measured under conditions of constant temperature, pressure, and solute chemical potential (ref 74).

Can k_D be accurately measured from sedimentation velocity experiments?

The short answer is no, it can't. First, remember that in a sedimentation velocity experiment it is the width and shape of the sedimenting boundaries that gives us information about diffusion. The reason k_D cannot be accurately measured is then that at solute concentrations high enough that the concentration dependence of D is significant, the width and shape of the sedimenting boundaries is profoundly affected (and often dominated) by the concentration dependence of the

sedimentation coefficient (as characterized by the k_s parameter). At high concentrations the concentration dependence of the sedimentation coefficient causes the sedimenting boundaries to exhibit so-called "self-sharpening", where they are much narrower than they are for the same material at low concentrations. This self-sharpening arises because the molecules sedimenting in the trailing portions of the boundary experience a low solute concentration, and therefore sediment significantly faster than the molecules in the plateau region and leading edge of the boundary, where the concentration is much higher. This effect means that the trailing edge of the boundary is constantly trying to catch up to the leading edge, leading to a narrower (sharper) boundary. That, in turn, makes it appear that the diffusion coefficient is much lower than it truly is.

Thus in a velocity experiment where solution non-ideality is significant, the effects of k_s on the boundary width and shape are typically stronger than the effects of k_D, making it quite difficult to separate these two effects. The only way to properly make this separation is to globally fit sedimentation velocity experiments run at multiple concentrations to a fitting model that accurately accounts for the non-ideality effects on both *s* and *D* (that is, a solution of the Lamm equation that includes the non-ideality effects, and where both k_s and k_D are fitting parameters), as was done for example by Wright *et al.* (2018). (Ref. 108)

In summary, if we plot the apparent $D_{20,w}$ from a sedimentation velocity experiment versus concentration, and then fit those data to a straight line, the slope of that line will in general not be the true k_D, but rather a value that depends strongly on the value for k_s. **Nonetheless, extrapolating that straight line to zero concentration will still give the correct value for $D^0_{20,w}$.**

Can the value for k_D be predicted?

By treating the non-ideality from a thermodynamic perspective, and by making an analogy between diffusion and osmotic pressure, it can be shown (Ref. 74) that *in the absence of reversible association*, and under conditions of constant temperature and constant solvent chemical potential, the value of k_D is directly related to the osmotic second virial coefficient and the concentration dependence of the sedimentation coefficient, through the relation

Equation 7:

$$k_D = 2BM - k_s$$

where *B* is the osmotic second virial coefficient (expressed in experimental units of mL mol/g rather than its molar counterpart B_{22}), *M* is the molar mass, and k_s is the coefficient for concentration dependence of the sedimentation coefficient (see Concentration Dependence of the Sedimentation Coefficient). Note that while equation 7 is often used in the literature for interpreting the "solution

interaction parameter", most such measurements are made using dynamic light scattering, where the condition of constant solvent chemical potential used in deriving equation 7 does not apply.

In addition to this thermodynamic approach, values for k_D can also be predicted from hydrodynamic theory or numerical simulations, where k_D can be directly related to the "swollen" specific volume, ν_s, which includes the bound or entrained solvent. From hydrodynamic theory Batchelor derived the following formula for (Ref. 71):

Equation 8:

$k_D = 8\nu_s - k_s$

It is worth noting that for an uncharged solute $2BM = 8\nu_s$, and therefore Eqs. 7 and 8 are equivalent for an isoelectric protein.

Continuing our derivations, k_s can then itself also be calculated as a multiple of ν_s (as discussed in more detail here). Using the preferred value (Ref. 74) of $k_s = 5\nu_s$ (from more recent Brownian dynamics simulations, ref. 72) we then get:

Equation 9:

$k_D = 8\nu_s - 5\nu_s = 3\nu_s$

Equations 8 and 9 do not, however, take into consideration the effect of the solute on the bulk viscosity of the solution. If we take the view that the viscosity "seen" by the diffusing solute is the actual solution viscosity rather than the viscosity of the pure solvent, the increased viscosity at high concentrations reduces the net increased diffusion rate, and thus the value for k_D will be substantially lower. This viewpoint makes sense for experiments where diffusion is driven by a macroscopic concentration gradient, but probably not for techniques where there is no concentration gradient and what is actually measured is the Brownian motion (*e.g.* dynamic light scattering or single particle tracking methods). For roughly spherical solutes it can be shown (ref. 74) that the increase in solution viscosity with increasing solute concentration nearly cancels the increase in diffusion coefficient calculated from Eq. 9, giving this formula for k_D :

Equation 10:

$k_D = 0.5\nu_s$

Overall then we see that Eq. 10 predicts a 6-fold smaller concentration dependence than Eq, 9, and also that k_D is 10-fold smaller than k_s. Does this make sense experimentally? Yes, the values predicted by Eq. 10 do appear to agree better with experimental data, at least for data obtained in experiments involving diffusing concentration gradients (ref. 74). On the other hand, there are many labs that use Eq. 7 to interpret their data and who argue it provides reasonable agreement with independently-determined values of *B*. There are also many reports of quite large values

for k_D; for example, Kingsbury et al. 2020 reported values from about -35 to +75 mL/g (ref. 75).

It should be noted that it has recently (2023-24) been shown that the swollen volume, hydration, and other hydrodynamic properties of proteins (Ref. 102), and even of flexible polymers such as PEG (Ref. 104), can be accurately calculated from their atomic models by a program called HullRad.

In conclusion, unfortunately different theoretical approaches and viewpoints yield quite different predicted values for k_D, and it remains unclear whether different types of diffusion experiments should be expected to give the same values. Consequently it is probably more important to actually make measurements at multiple concentrations and then do a real extrapolation to c = 0 for diffusion than it is for sedimentation.

What is the rationale for the initial default value for k_D?

The default value for k_D is 0.9 mL/g when the program is installed, which is 1/10 of the default value for k_s , as predicted by Equation 10. That default value then essentially represents the minimum concentration dependence for D one could expect to see, and based on the arguments and experimental data discussed in reference 74 this value probably makes sense for experiments involving diffusing concentration gradients (but might not really make sense for data from dynamic light scattering).

Definitely what (if any) default value should be used for k_D is much more uncertain than any of the other SEDNTERP default values. If the user prefers, it may be reasonable to make the default k_D be consistent with Eq. 9 and assuming $k_s = 5\nu_s$ (ref. 72), which makes the default k_D 0.6 times the default k_s.

3.1.8 The Svedberg Equation

One key equation for AUC interpretation is the Svedberg equation (Ref. 68), which relates the molar mass to the sedimentation coefficient, s, and the diffusion coefficient, D:

Equation 2:

$$M = \frac{sRT}{D\left(1-\bar{v}\rho\right)}$$

where R is the gas constant and T is the temperature (Kelvin).

The Svedberg equation shows that molar mass can be determined from the ratio of the

sedimentation to diffusion constant. It also shows that the three important quantities *s*, *D*, and *M* are inter-dependent: each can be calculated if the other two are known. That fact is often used, for example, in modern data analysis packages for sedimentation velocity data to allow fitting the molar mass of each species rather than its diffusion coefficient (since the analyst is often more interested in knowing *M* rather than *D*). This is done by internally using Eq. 2 to calculate the *D* value needed to characterize the boundary spreading from the current *s* and *M* values.

3.1.9 Limitations

There are four categories of limitations in the calculations implemented in SEDNTERP:

Limitations implied in choosing a method of interpretation.

SEDNTERP implements certain general models for interpretation which are not always the best models for a certain experiment. This kind of limitation can be overcome by extending the principles used in SEDNTERP to include more sophisticated algorithms. In these cases SEDNTERP may still be useful to compute some parameters needed in these more sophisticated algorithms. Users who often use an algorithm not implemented in SEDNTERP should contact the programmers; the source code is freely available, and suggestions for extensions of general use are welcome. For example, the SEDNTERP program assumes that only sedimentation or diffusion data are available, and calculations or methods that rely on knowledge of the viscosity increment (Ref. 31) or the radius of gyration (Ref. 32) are not implemented and will not be discussed here. As another example, modeling proteins as ellipsoids of revolution is most applicable to the examination of monomeric, globular proteins, and is less useful in the detailed examination of oligomeric structures. To overcome this limitation, either more exact or more general models should be used. (Ref. 58) However, the use of such models requires that the experimenter already have some structural information available. In these cases, provided that sufficient material of high enough purity is available, there are other experimental methods, such as light, x-ray, or neutron scattering, which can provide more detailed structural information. That is to say, the real power of analytical ultracentrifugation is to provide an initial, rough idea of the asymmetry or hydration of a protein. (Refs. 56, 57) The field of hydrodynamic modeling and interpretation, and the combination of AUC methods with structural tools such as x-ray scattering, is quite an active one, and the bibliography in this document is far from current in that area.

Incomplete separation of information

A second limitation is that information that can be derived from the experimental data about certain molecular properties is not cleanly separated from information about other properties (the properties are correlated or "not orthogonal"). That is particularly true for hydration and

asymmetry, and that limitation is discussed in a separate topic.

Limitations fundamental to the use of predictive schemes

Many of the equations are derived assuming a two-component solution, even though experiments routinely are performed in buffers with three or more components. On theoretical grounds it is clear that this two-component assumption is not correct. Nonetheless, over a reasonable protein concentration range, and when thorough dialysis has been performed against solvents of moderate pH (i.e. where the net-charge to molar mass ratio of the protein is less than about 1:5000 mole/g) and moderate ionic strength (0.05-0.5 M), the sedimentation behavior of many proteins is described adequately by the methods used here (Refs. 2, 3). It is still important that the experimenter be aware that neglect of complexities resulting from multi-component solutions can lead to misinterpretation of data (Refs. 3, 13).

Limited accuracy of interpolated values

The interpolating functions used by SEDNTERP are convenient and quick, but may not give sufficient accuracy to suit the user's purpose. Without compelling reason, none of the calculated or predicted values should be used as a substitute for real experimental determinations. For example, the molar mass calculated from $\bar{\nu}$, ρ, ω^2, and σ should not be used as a substitute for the molar mass calculated from a known composition, unless there is a compelling reason such as post-translational modification of the protein. Likewise, the values of $\bar{\nu}$, ρ, and η computed by SEDNTERP should never be used as a substitute for experimentally determined values. For example, due to non-additive volume changes on mixing, densities calculated by SEDNTERP using equation 13 may be in serious error when two or more solutes are present in moderate concentrations. This same problem can arise for mixtures containing both salts and a high mole fraction of non-ionic components. In fact, SEDNTERP will compute the density and viscosity of buffers that could never be created (due to solubility limits). Note that for every physical parameter that SEDNTERP interpolates or estimates an entry field for exact experimental data is included.

3.2 Key Molecular Properties

3.2.1 s*, Apparent Sedimentation Coefficient

The "apparent" or "raw" sedimentation coefficient, *s** (pronounced 'ess star'), is the actual experimental value measured by sedimentation velocity under particular experimental conditions

(temperature, buffer density and viscosity, and solute concentration). Thus it is "apparent" in the sense that the value is condition-dependent.

Like all types of sedimentation coefficients, the units of s* are seconds. The value of *s** is the measured speed of the second moment (or as a good approximation, the peak of the derivative) of a boundary in a sedimentation velocity experiment. *s** is not corrected for density, viscosity, or solute concentration. Usually, *s** is also not corrected to account for heterogeneity or other forms of non-ideality in the sedimenting boundary.

Having a value for either the sedimentation coefficient or the diffusion coefficient is required in order to obtain information about the molecule's hydration and asymmetry. Before such interpretation is possible, it is also necessary to correct the s* value to standard conditions, which requires information about the partial specific volume, buffer density, and buffer viscosity.

Caution: Do not confuse the *s** symbol or the term "apparent sedimentation coefficient" as used in this document with the same symbol and terminology as used in the context of the sedimentation coefficient distributions *g*(*s**), *ls-g*(*s**), or *G*(*s**). In that context the "apparent" adjective and the '*' symbol has to do with spreading of sedimentation boundaries by diffusion, not to the fact that the value depends on buffer density, viscosity, or temperature.

3.2.2 D*, Apparent Diffusion Coefficient

The "apparent" or "raw" diffusion coefficient, *D** (pronounced 'dee star'), is the actual experimental value measured under particular conditions (temperature, buffer viscosity, and solute concentration). Thus *D** is "apparent" in the sense that the value is condition-dependent.

To measure the diffusion coefficient, one can (1) directly monitor the spreading of a concentration gradient over time in synthetic boundary cells at low rotor speed in an AUC instrument (Ref. 108), or (2) fit sedimentation velocity data using models that account for the effects of diffusion on the width and shape of the boundaries (Refs. 109, 110), or (3) less directly monitor a spreading concentration gradient via Taylor dispersion (Ref. 111).

Alternatively, *D** can be measured in the absence of a concentration gradient, but indirectly, by methods that detect the Brownian motion of the macromolecule, such as dynamic light scattering (Ref. 112) or single-particle tracking approaches (Ref. 113).

Like all types of diffusion coefficients, the units of D* are cm^2/sec, but are commonly (and more conveniently) expressed in Ficks (symbol F), where 1 Fick = 1×10^{-7} sec. The value of D* has not been corrected for viscosity, temperature, or solute concentration.

Having a value for either the diffusion coefficient or the sedimentation coefficient is required in order to obtain information about the molecule's hydration and asymmetry. Before such interpretation is possible, it is also necessary to correct the D* value to standard conditions, which requires information about the buffer viscosity and the experimental temperature.

3.2.3 Partial Specific Volume

The partial specific volume, which is usually given the symbol $\bar{\nu}$ (pronounced "vee-bar"), has units of milliliters per gram (mL/g). Thus, to a first approximation, it is simply the inverse of the density. However, partial specific volume is more strictly defined as the increase in volume when 1 gram of protein is mixed into a solution, and that volume change includes the effects the protein has on the volume of the solvent near it. Therefore a protein may have different partial specific volumes for different buffers. Some substancesmay even have a negative partial specific volume (like ammonium chloride in water) because instead of taking up volume upon mixing, they actually reduce the volume of the solution by condensing water around themselves.

Why is $\bar{\nu}$ a key parameter?

The partial specific volume is important for AUC experiments because it determines the buoyancy force which opposes the centrifugal force applied by the rotor, and it is only the net force (centrifugal minus buoyancy) that drives sedimentation. Not only does the value of $\bar{\nu}$ enter directly into computing the buoyant molar mass, it also provides one measure of the volume the macromolecule occupies in solution, which is important for determining hydrodynamic properties such as the molecule's minimum frictional coefficient, f_0, (and therefore its maximum sedimentation coefficient). The partial specific volume is also used as the molecular volume when estimating the specific refractive increment, *dn/dc*.

For typical proteins, any fractional errors in $\bar{\nu}$ are magnified ~3-fold in the determination of molar mass from AUC data (Ref. 9). Therefore, good estimates of $\bar{\nu}$ are necessary for accurate interpretation of AUC data. An excellent discussion of partial specific volumes, along with an extensive tabulation of values, may be found in the review by Durchschlag (Ref. 6). Several methods are available for the measurement of $\bar{\nu}$, but these typically require relatively high amounts and concentrations of solute, usually require specialized equipment, and definitely require careful analytical techniques for their success (Refs. 6, 40). Accurately knowing the protein solution's weight concentration is often another limiting factor for accurately measuring $\bar{\nu}$.

An alternative to measuring $\bar{\nu}$ is to estimate its value based on the protein's composition. Comparison of measured and estimated $\bar{\nu}$ values indicate that the error typically is less than 1% (Ref. 40). Somewhat larger errors are observed for small peptides, proteins that have an unusual amino acid composition, proteins that are conjugated with carbohydrate, proteins that exhibit preferential hydration, or proteins that have a non-globular structure (Refs. 6, 40, 41). Even under these circumstances, the error is typically less than 3% if the correct composition is used (Refs. 6, 40).

3.2.4 Buoyant Molar Mass

The **buoyant molar mass**, M_b, of a molecule is its true molar mass, M, less the mass of solvent (buffer) it displaces. In an AUC experiment it is the **buoyant** molar mass, not the total (true) molar mass, that provides the driving force for sedimentation.

Note that unlike the molar mass, the buoyant molar mass is not an intrinsic property of the molecule, it is a property that depends on experimental conditions (specifically the density of the buffer and the partial specific volume).

Also note that for molecules with very low density (such as lipids or detergents) the buoyant molar mass can actually be **negative** (in the centrifuge such molecules float toward the center of the rotor rather than sediment toward the outside).

3.2.5 Hydration and Asymmetry

The sedimentation and diffusion rates of a macromolecule are limited by hydrodynamic friction, which is measured by the frictional coefficient, f. That frictional coefficient depends on the shape (conformation) of the molecule: it will be much larger for a highly extended, rod-like shape, or for a random coil, relative to a compact, globular shape. The experimental value for f can be obtained from the sedimentation coefficient using Eq. 4, or from the diffusion coefficient using Eq. 3.

Because the experimental f depends on buffer viscosity, before trying to interpret f (for example, comparing values between molecules) it is desirable to correct the value to a standard condition, where the viscosity is the viscosity of water at 20 °C. Thus the standardized f can be calculated simply by inserting the $s^0_{20,w}$ or $D^0_{20,w}$ value into Eqs. 4 or 3, respectively. SEDNTERP always calculates and displays this standardized value for f, which could (and probably should) be labeled as an "$f_{20,w}$" value, but such nomenclature is not the common practice in this field.

The overall hydrodynamic compactness of a molecule is then often characterized by reporting the ratio f/f_0, where f_0 is the minimum possible frictional coefficient for this molecule (the value for an anhydrous sphere of equivalent volume, see Eq. 43). An f/f_0 ratio of 1.2-1.3 usually indicates a fairly compact structure with low asymmetry and disorder, whereas a ratio above 1.6 suggests an extended structure with fairly high asymmetry (or perhaps a partially-disordered protein). Thus either a sedimentation or diffusion coefficient can, in principle, provide valuable information about the shape/conformation of macromolecules in solution. Often that shape information is interpreted in terms of degree of molecular asymmetry using simple shape models such as ellipsoids or rods, with the degree of asymmetry quantified by the axial ratio or length/diameter ratio of such a model.

However, hydration also contributes directly to the friction (since the bound or entrained water moves along with the macromolecule). The degree of hydration of proteins is usually quantified as grams of bound water per gram of protein, and is given the symbol δ_1. Based on sedimentation or diffusion coefficients alone it is simply not possible to uniquely separate the contributions of hydration *vs.* shape/conformation to f (Refs. 28, 29, 30). In the absence of other external information about hydration or shape it is only possible to calculate a *maximum* degree of hydration or a *maximum* asymmetry parameter (*e.g.* axial ratio).

SEDNTERP calculates these estimates of hydration or shape using the models of prolate ellipsoid, oblate ellipsoid, rod, or disk for the asymmetry. Since hydration and asymmetry both affect the friction, only maximum values for hydration or asymmetry can be given from sedimentation experiments alone. However, if some information is available about either hydration or asymmetry, then SEDNTERP will use that information to calculate better estimates of the other value. For instance, if the protein composition is used for an estimate of hydration, you may optionally use this estimated hydration to calculate a better estimate of the asymmetry.

3.3 Calculations Implemented in SEDNTERP

3.3.1 Calculating Molar Mass

SEDNTERP can use 4 different options to determine which molar mass value will be used in subsequent calculations:

1. Molar mass from direct input

With this option, the program does not calculate the molar mass at all, but simply uses the value entered by the user. Molar mass estimates obtained from mass spectroscopy, sequence analysis,

gel electrophoresis, gel filtration or other methods may be entered explicitly by clicking on the "Use known value" button within the Molar mass frame on the Sample Properties major tab. The text box containing the molar mass will get the focus and you may type in a molar mass value (grams/mole).

NOTE: The molar mass needed for proper interpretation of AUC data is that for the protein in solution, which is not necessarily the same as that measured by mass spectroscopy (or calculated by mass spectroscopy software), where the protein is an ion, where it may have had its disulfides reduced, and where oligomers stabilized by non-covalent interactions will usually be dissociated.

2. Molar mass from composition

SEDNTERP can calculate a molar mass based upon composition, by summing the molar masses of each of the protein's components. This calculation is performed in the Sample Composition Sample Composition form available by pressing the "From Composition" button within the Molar mass frame. The *minimum* molar mass is estimated by summing up the g/mole of each of the protein's components.

The amino acid composition can be entered as mole-residue per mole-protein in the grid control, with the residues sorted by either one-letter or three-letter amino acid codes. Entries can also be automatically tabulated from amino acid sequences pasted from the Windows Clipboard or entered from the keyboard. On a separate tab are entries for common protein ligands ("conjugates").

CAUTION: It is important to note that the molar masses for the amino acids have been adjusted to reflect the loss of H_2O associated with the formation of amide linkages. (The program adds the weight of one additional H_2O per polypeptide chain since there is one less linkage than the total number of amino acids.) However components in the conjugates table have *not* have been corrected in this fashion, so be careful. For complex carbohydrates there is an entry for a 'glycosidic linkage' you can use which has a *negative* molar mass. You may browse the values given for the amino acids and conjugates in the program database from the Estimating Databases Menu.

It must be realized that the molar mass calculated in this fashion can be in serious error if the composition is based on DNA sequence data and significant post-translational modification of the protein occurs or if the protein exists as an oligomer. If the protein is *known* to be an oligomer, that fact and its known stoichiometry can be specified on the composition form and then all molecular properties will be adjusted appropriately.

NOTE: The molar mass the SEDNTERP computes for polypeptides does not reflect the small changes associated with loss or gain of hydrogen atoms by the ionizable groups as a function of pH. That is, the sequence mass (as well the partial specific volume and *dn/dc*) are assumed to be independent of pH.

Given a composition, other properties of the sample also are automatically calculated, including the estimated hydration, the partial specific volume, the charge, the isoelectric point, extinction coefficients, extinction spectrum, and *dn/dc*.

3. Molar mass from sedimentation equilibrium

SEDNTERP can use molar mass values determined by analyzing sedimentation equilibrium data, as entered onto the Sedimentation Equilibrium minor tab. Sedimentation equilibrium results may be returned in different units, depending on which data analysis package was used, and depending on whether information about the sample $\bar{v}$ and the buffer density were available at the time of data analysis. Therefore sedimentation equilibrium results can be entered in one of three different types of units:

1. As the true molar mass, M_r
2. As the buoyant molar mass, M_b
3. As the quantity σ (the 'reduced apparent molar mass')

When σ units are used, the true molar mass is calculated using the formula (a form of Eq. 14):

$$M = \frac{\sigma RT}{\left(1-\bar{v}\rho\right)\omega^2}$$

Obviously using σ units requires that accurate values for $\bar{v}$, ρ, and the rotor speed are known.

When buoyant mass units are used, then the true molar mass *M* is calculated using the formula (Eq. 1)

$$M_b = M\left(1 - \bar{v}\rho\right)$$

which obviously requires accurate values for $\bar{v}$ and ρ.

4. Molar mass from an s/D ratio

The molar mass can be also calculated from the ratio of sedimentation to diffusion coefficients using the Svedberg equation (Equation 2):

$$M = \frac{sRT}{D\left(1 - \bar{v}\rho\right)}$$

This *s/D* ratio could come directly from a single sedimentation velocity experiment where the data analysis returned both *s* and *D* (or *s* and *M*), or it could come from separate sedimentation and diffusion experiments.

NOTE: Modern sedimentation velocity data analysis packages often fit a specific sedimenting boundary and return its sedimentation coefficient and molar mass. However *internally* such packages are really fitting *D*, not *M*, but then internally using the Svedberg equation to report *M* rather than *D*. Thus SEDNTERP logically considers such a result to be 'obtaining *M* from the *s/D* ratio', even though *D* may not have been explicitly reported.

3.3.2 Calculating Buoyant Molar Mass

SEDNTERP will calculate the buoyant molar mass of the sample according to the method chosen for calculating the molar mass.

The buoyant molar mass M_b is defined as the true molar mass, *M*, less the mass of solvent (buffer) the macromolecule displaces. It is calculated via:

Equation 1.

$$M_b = M\left(1 - \bar{v}\rho\right)$$

where *M* is the true molar mass, $\bar{v}$ is the partial specific volume, and ρ is the buffer density (not the density of the actual protein solution).

If the molar mass is computed from σ, then the buoyant molar mass is computed from σ also, via this alternate form of Equation 14:

$$M_b = \frac{\sigma RT}{\omega^2}$$

If the molar mass is entered directly, or if it is calculated from composition, the buoyant molar mass is computed from Equation 1 above using the values for $\bar{v}$ and ρ from the current buffer.

3.3.3 Calculating Vbar Estimates

Due to the difficulty of accurately measuring , it is common practice to instead estimate (calculate) based on the protein's composition. Such estimates have typically (Refs . 5, 6, 40) been calculated by assuming the volumes of the macromolecule's constituents are simply additive, using the method of Cohn and Edsall (Cohn and Edsall, 1943) (Ref. 4):

Equation 15:

$$\bar{v} = \frac{\sum n_i M_i \bar{v_i}}{\sum n_i M_i}$$

where *n* is the number of moles of the i^{th} component and M_i is its molar mass. Values of $\bar{v}$ for the amino acids and several protein-associated ligands ("conjugates") are stored in the program database. Note that traditionally the M_i values used are the *residue* weights, *i.e.* they have been adjusted to reflect the loss of H_2O associated with the formation of the amide linkage. This means that this formula essentially ignores both the molar mass and the molar volume of the amino-terminal proton and the carboxy-terminal hydroxyl (and this was the implementation used in SEDNTERP versions 1 and 2).

However in SEDNTERP version 3 it is possible to include two additional terms that were dropped in the Cohn & Edsall procedure, but which should improve the accuracy (especially for small proteins and peptides). One term accounts for the mass and volume of the terminal groups in the calculations, and the other is a correction for the Traube "covolume" (Refs. 94, 95), an empirical term which adds extra volume (once per molecule) to account for imperfect packing of the atoms. With these two additions Eq. 15 becomes

Equation 16:

$$\bar{v} = \frac{\sum n_i M_i \bar{v}_i + n_{chains} V_{term} + V_{cov}}{\sum n_i M_i + n_{chains} M_{water}}$$

where n_{chains} is the number of separate amino acid chains represented in the current amino acid composition, V_{term} is the total molar volume of an N- and C-terminus (in mL), V_{cov} is the Traube covolume (in mL), and M_{water} is the total molar mass of a pair of terminal groups. The values of V_{term} and V_{cov} are quite small compared to the total volume of the residues for most proteins, and thus for a 50 kDa protein these additional terms change the calculated $\bar{v}$ only by a few digits in the 4th decimal place (consistent with Cohn & Edsall's conclusion that they are negligible and can be dropped). However the importance of these terms is inversely proportional to the total molar mass, so these two terms will be much more significant for small peptides.

NOTE: While the use of these additional terms is recommended, each of them is optional, and can be turned on or off using check boxes found on the Sample Composition form. When the terminal groups are added to the $\bar{v}$ calculation, for consistency they are also included in the *dn/dc* calculation (which depends on $\bar{v}$).

Whether each of these terms is normally added for new sample compositions is controlled by separate program default settings (alterable via the **Alter program defaults...** item on the File Menu). Note also that when Sample records are imported from old version 1 or version 2 databases these extra terms in Eq. 16 are not used (any $\bar{v}$ value previously calculated using Eq. 15 will not be automatically re-calculated when the record is imported).

The default value used for the V_{cov} when the program is installed is 12.4 mL/mole (ref. 94). The default value for V_{term} is -6.4 mL/mole (as kindly calculated using the methods of Ref. 94 by Mattia Rocco). Note that this value is negative because the two terminal groups are assumed to be charged, and the resulting electrostriction actually exceeds the group volumes. Both of these recommended values can be altered via the **Alter program defaults...** item on the File Menu should the user desire to use different values.

TIP: The conjugates list includes new items to help provide $\bar{v}$ estimates for glycoproteins and PEGylated proteins. The 'N-linked glycans' item represents 1 kDa of glycan with an average $\bar{v}$ value from the literature (taken from ref. 80). Thus if your glycoprotein on average contains 5.4 kDa of glycans, you would add 5.4 moles of this conjugate. Similarly the 'polyethylene glycol 10 kDa' item has a $\bar{v}$ value taken from ref. 81, and can be used to add the appropriate mass of PEG.

Comparison of measured and estimated $\bar{v}$ values indicate that the error in such $\bar{v}$ estimates typically is less than 1% (Ref. 40). Somewhat larger errors are observed for proteins that have an unusual amino acid composition, proteins that exhibit preferential hydration, proteins that have a non-globular structure, and (using the Cohn and Edsall procedure) for small peptides, but even in such cases the error is typically under 3% (Refs. 6, 40, 41). For glycoproteins where the carbohydrates represent a significant portion of the total mass it will of course be important to account for the $\bar{v}$ of the carbohydrate portion (Refs. 40, 80)

Correcting for temperature

SEDNTERP uses the values in the database to calculate $\bar{v}$ at 25 ^{o}C and then uses **equation 17** to adjust this value over a temperature range of 4-45 ^{o}C (Ref. 6):

Equation 17:

$$\bar{v}_T = \bar{v}_{25} + 4.25 \times 10^{-4} (T - 25)$$

[note that this equation is incorrect in ref. 59]. This same temperature correction is also applied to any added conjugates.

Other potential corrections

There is no computational method to correct for the effects of pH. However, it has been observed that $\bar{v}$ varies only slightly with pH unless denaturation occurs, at which point $\bar{v}$ decreases by about 3%. This decrease in $\bar{v}$ with pH parallels other changes in the protein structure (e.g. unfolding as monitored by viscosity or by circular dichroism).

Buffer composition and preferential solvation also can effect $\bar{v}$ (Ref. 6), but these are usually neglected except for denaturing buffers.

All of the partial specific volumes in the program database have been determined for unconjugated molecules. For a few non-polar ligands, this approximation has been shown to be reasonable, however, caution should be exercised when calculating partial specific volumes for lipoproteins (Ref. 35). Likewise, greater uncertainty results from the use of $\bar{v}$ determined for unconjugated carbohydrates (Ref. 41).

Calculating denatured vbar estimates

It can be useful or necessary to measure the molar mass of proteins in denaturants to dissociate non-covalent oligomers into monomers. Such an AUC experiment used to be routinely done for each new protein, but since we now nearly always know the sequence mass of our protein, this is not needed nearly as often as it was prior to the 1990's.

For denatured proteins, preferential binding of the denaturant to the protein will affect the apparent partial specific volume and must be taken into account. That is, we want to measure the mass of the protein **without** the bound denaturant, and we must adjust the $\bar{v}$ value in order to do that. Under these circumstances, $\bar{v}$ must be replaced by the apparent isopotential partial specific volume (Ref. 42) ϕ_2' (the nomenclature here follows that of Durchschlag, Ref. 6, where component 2 is the protein). (Refs. 3, 6) In general, increasing ionic strength results in a linear increase in ϕ_2', with the slopes following the Hofmeister series. However, preferential hydration can lead to nonlinear changes in ϕ_2'. Values of ϕ_2' can be determined as described by Lee and Timasheff (Ref. 42) and then entered in directly using the "Use known value" button.

However, special provisions are made in SEDNTERP for estimating ϕ_2' of simple proteins in solvents containing either 8 M urea or 6 M guanidinium chloride. Using methods described in Refs. 6, 37, 38, and 42, ϕ_2' can be calculated via

Equation 18:

$$\phi_2' = \bar{v} - \left\{\frac{1}{\rho} - \bar{v}_3\right\}\{\delta_3 - g_3\delta_1\}$$

where $\bar{v}_3$ is the partial specific volume of the denaturant, δ_3 is the number of grams of denaturant bound per gram of protein, δ_1 is the hydration in grams of water per gram of protein, and g_3 is the number of grams of denaturant per gram of water. (Note that the sign of the second term in equation 18 is incorrect in ref. 59.]

The ϕ_2' determined by equation 18 (ref. 37) then replaces $\bar{v}$ in all calculations performed by

SEDNTERP.

For 6 M guanidine HCl at 20 ^{o}C, the values are ρ = 1.1418 g/mL, $\bar{v}_3$ = 0.763 mL/g and g_3 =1.007 g guanidine-HCl per g H_2O (Ref. 42). For 8 M urea at 20 ^{o}C the corresponding values are ρ = 1.1152 g/mL, $\bar{v}_3$ = 0.763 mL/g and g_3 = 0.752 g urea per gram of water (Ref. 37).

For both urea and guanidine HCl, δ_3 is calculated by assuming one molecule of denaturant is bound per pair of peptide bonds and one to each aromatic side chain, including histidine (Refs. 37, 42), giving the following equation:

Equation 19:

$$\delta_3 = \frac{M_d}{M_p}\left(\frac{N_t - 1}{2} + N_{aromatic}\right)$$

where N_t is the total number of amino acids, $N_{aromatic}$ is the number of aromatic residues, M_d is the molar mass of the denaturant (95.54 for guanidine and 60 for urea) and M_p is the subunit molar mass for the protein.

Note that erroneous values of ϕ_2' may be calculated if the overall protein composition has significant levels (> ~10% wt/wt) of non-amino acid constituents (Ref. 43). However, empirical evidence suggests that reasonable estimates for ϕ_2' are obtained for glycoproteins in 6 M guanidine HCl by assuming each carbohydrate moiety binds one guanidine molecule and by assuming δ_1 is not affected by the presence of the carbohydrate (Ref. 43). In this case, equation 19 is modified to account for the guanidine binding, becoming:

Equation 20:

$$\delta_3 = \frac{M_d}{M_p}\left(\frac{N_t - 1}{2} + N_{aromatic} + N_{carbohydrate}\right)$$

There has not been sufficient systematic experimental evidence to justify the use of a similar analysis for 8 M urea denaturation of glycoproteins. (Ref. 37)

Non-denaturants may benefit from similar analysis. Timasheff and co-workers have described calculations of ϕ_2' for simple proteins in solvents containing varying amounts of NaCl, Na_2SO_4, $MgSO_4$, glycine, -alanine, betaine and CH_3COONa (Ref. 38). It should be noted that such calculations require knowledge of the interaction (i.e. solvation) parameters (δ_1 and δ_3) for the

particular salt, and that it is incorrect to attempt these calculations using interaction data for a different salt (Ref. 38).

3.3.4 Interpolating a Buffer Density

To interpolate a density for a buffer press the **From Composition** button within the Density panel on the Buffer Properties major tab. The Buffer Composition form will then appear. This same form will also allow you to calculate the buffer viscosity and refractive index, and optionally use those calculated property values as the properties of your buffer.

TIP: Measuring a buffer density with sufficient precision to get an accurate molar mass or $s_{20,w}$ value does not require expensive, specialized equipment (unlike measuring a $\bar{v}$). All that is needed to measure the specific gravity of your buffer is a good laboratory balance and a specific gravity bottle (also known as a pycnometer), which are available from major lab glassware suppliers in volumes from 5 mL and up for ~$100 or less.

The first step SEDNTERP uses to calculate the density of a buffer is to calculate the density of water at 20 ^{o}C (including the effects of any isotopic variants of water that are present in the buffer. Then the density of the buffer at 20 ^{o}C can be calculated based on its composition using 'density increments'. Finally, the density is corrected for temperature assuming that water (or the appropriate mixture of water isotopes) is the predominant component in the buffer. While that assumption is true for solutions containing moderate amounts of other components, significant errors may be introduced for solutions containing high solute concentrations (Ref. 8).

Calculating the temperature dependence of water density

In order to correct a buffer density known at one temperature to any other temperature, a fundamental assumption is made that the presence of the buffer solutes does not alter the temperature dependence of the density of water (or its heavy isotopes). This should be true for buffers containing moderate concentrations of other components, but significant errors will potentially be introduced for high solute concentrations.

To calculate the density of water as a function of temperature at 1 atm of pressure, SEDNTERP version 3 uses the approach and equations recommended by The International Association for the Properties of Water and Steam (ref. 76). First, the specific volume of water v at the reference pressure p_0 = 0.1 MPa and temperature T is calculated using this 12-term polynomial:

Equation 21:

$$v(p_0,T)=\frac{RT_R}{p_0}\left\{a_1+\sum_{i=2}^{6}a_i\alpha^{n_i}+\sum_{i=1}^{6}b_i\beta^{m_i}\right\}$$

where R = 461.51805 J·kg^{-1}·K^{-1}, T_R = 10 K, and the dimensionless coefficients a_i and b_i and the exponents n_i and m_i, are listed in the table at the end of this topic. The dimensionless parameters α and β in Eq. 21 are given by:

Equation 22:

$$\alpha=\frac{T_R}{T_a-T};\quad \beta=\frac{T_R}{T-T_b}$$

where T_a = 593 K and T_b = 232 K.

Next the isothermal derivative of specific volume with respect to pressure, $v_p=\left(\partial v/\partial p\right)_T$, is calculated at temperature T and p_0 = 0.1 MPa using this 12-term polynomial:

Equation 23:

$$v_p(p_0,T)=\frac{RT_R}{{p_0}^2}\left\{\sum_{i=7}^{11}a_i\alpha^{n_i}+\sum_{i=7}^{13}b_i\beta^{m_i}\right\}$$

where again the coefficients a_i and b_i, and the exponents n_i and m_i, are given in the table at the end of this topic. That isothermal derivative value is then used to correct the specific volume to 1 atm pressure (0.101325 MPa), and the desired density is then the inverse of this final specific volume. Densities thus calculated are accurate to 1 ppm over the 0-100 ^{o}C range (ref. 76).

Allowing for isotopic variants of water

To handle cases where the buffer contains heavy isotopes of water (D_2O, H_2O^{18}, or D_2O^{18}) SEDNTERP uses the methods described by Steckel and Szapiro (1963) (Ref. 45). The density of each isotopic form of water can be computed using

Equation 24:

$$\rho_T = \rho_{\max}\left\{1 - \mathrm{F}\left(T - T_{\max}\right)^2\left[1.74224 + \frac{482.502}{\left(T - T_{\max}\right) + 77.861}\right]\right\} \times 10^{-6}$$

where ρ_T is the density at temperature T in C, $\rho_{\max}$ is the maximum density (observed at temperature $T_{\max}$), and F is an empirical constant. Equation 12 may be used for all water species by substituting in the appropriate values from the following table. The values of $\rho_{\max}$, $T_{\max}$, and F for each heavy isotope are listed in the table immediately below. Equation 24 is valid from the melting temperature to about 80 °C.

Isotope	$T_{\max}$	$\rho_{\max}$	F	melting T
H_2O^{18}	4.30	1.11255	1.0	0.28
D_2O	11.24	1.10602	1.0555	3.81
D_2O^{18}	11.46	1.21681	1.0555	4.02

It should be noted that the values in the table above reflect natural isotopic abundance for the O and H in D_2O and H_2O^{18}, respectively. Similarly, the density for H_2O derived via Eqs. 21-23 is for water at natural isotopic abundance. The density for mixtures of different isotopic forms of water are calculated assuming there is no volume change upon mixing (Ref. 45), which means the density for each isotopic form is weighted according to its volume fraction.

Estimating the density from the buffer composition

The first step SEDNTERP uses to estimate the density of a buffer is to calculate the density of the buffer at 20 ^{o}C using the concept of 'density increments'. The density increment for each solute is the density difference between a solution containing only that solute and pure water. These density increments are then summed to give the total difference between the buffer and pure water. That is, the assumption is that each solute's effect on the density is independent from the others.

To determine the density increments, tables of the density of solutions of 102 different single solutes *vs.* molar concentration were fitted by least squares methods to a common polynomial form. The sets of best-fit polynomial coefficients for each solute are then stored in the program

database. From this, the density of a single solute solution anywhere within the original fit concentration range can be interpolated. The density increment for the *i*th buffer component at molar concentration C_i is then calculated as:

Equation 25:

$$\Delta\rho_{C_i} = a_i + b_i \times 10^{-3}\sqrt{C_i} + c_i \times 10^{-2} C_i + d_i \times 10^{-3} C_i^2 + e_i \times 10^{-4} C_i^3 + f_i \times 10^{-6} C_i^4 - \rho_{water}$$

where *a*-*f* are the coefficients from the polynomial fit using least squares methods (Ref. 46) and ρ_{water} is the density of pure water.

The inclusion of the square root term improves the fit at low concentrations for some ionic solutes. In principle, *a* should be the density of pure water at 20 ^{O}C (Ref. 44), but substantially improved fits are sometimes obtained when this term is allowed to vary. Thus, equation 25 is only an empirical fitting function, and, as such, no physical interpretation of the parameters should be made. For each solute only the polynomial terms required to obtain an adequate fit (no systematic errors as a function of concentration) were retained (that is, one or more of the coefficients *b*-*f* were held fixed at zero, or *a* was held fixed at the density of water).

The concentration range over which each set of coefficients is valid, and the maximum error of the fit over that concentration range, are recorded in the program database. In general, polynomials fit the density data for simple salts far better than for organic salts and alcohols. Systematic deviations are observed in fits for alcohols, with that for ethanol being the worst case. Even though the magnitude of the errors is within the last decimal place of the tabulated data, it is clear that the polynomial function provides only a marginally adequate model for those substances. For most components, the precision of the fit using SEDNTERP is better than 1×10^{-4} g/mL. For a buffer density » 1 and a typical protein $\bar{v}$ of 0.73 mg/ml, an error of that magnitude in the density will produce only an 0.03% error in the calculated molar mass, which is far smaller than the likely experimental precision.

Having computed the density increments for each component, SEDNTERP will calculate the density of the buffer at 20 ^{O}C (including the effects of any heavy isotopes of water) by simply adding the density increments to the density of the current isotopic mixture of water, ρ_{heavy}, at 20 ^{O}C:

Equation 26:

$$\rho_{buffer,20} = \rho_{heavy,20} + \sum_i \Delta\rho_{C_i}$$

Finally, the buffer density at the desired temperature *T* is calculated using:

Equation 27:

$$\rho_{\text{buffer},T} = \rho_{\text{buffer},20} \times \frac{\rho_{\text{heavy},T}}{\rho_{\text{heavy},20}}$$

where both $\rho_{\text{heavy,T}}$ and $\rho_{\text{heavy,20}}$ include the effects of any heavy isotopes of water in this buffer.

Coefficients and exponents for Equations 21 and 23.

i	n_i	m_i	a_i	β_i
1	-	1	1.93763157E−2	5.78545292E−3
2	4	2	6.74458446E+3	−1.53195665E−2
3	5	3	−2.22521604E+5	3.11337859E−2
4	7	4	1.00231247E+8	−4.23546241E−2
5	8	5	−1.63552118E+9	3.38713507E−2
6	9	6	8.32299658E+9	−1.19946761E−2
7	1	1	−7.5245878E−6	−3.1091470E−6
8	3	3	−1.3767418E−2	2.8964919E−5
9	5	4	1.0627293E+1	−1.3112763E−4
10	6	5	−2.0457795E+2	3.0410453E−4
11	7	6	1.2037414E+3	−3.9034594E−4

12	-	7	-	2.3403117E−4
13	-	9	-	−4.8510101E−5

3.3.5 Interpolating a Buffer Viscosity

Viscosity is a measure of the energy dissipated in a fluid in order to maintain flow. The viscosity of the experimental buffer will affect diffusion coefficients, sedimentation coefficients, and also the time needed to reach sedimentation equilibrium (Refs. 2, 14, 17). Excellent reviews are available describing methods for determining viscosity and for interpreting viscosity data (Refs. 47, 48). The buffer viscosity is dependent on the buffer composition, and also **very** strongly dependent on temperature. Accurate buffer viscosity values are required to accurately standardize sedimentation and diffusion coefficients. Unfortunately however viscosity instrumentation with precision and accuracy comparable to that of the raw sedimentation and diffusion coefficients from current instruments is unfortunately not available in most biochemistry and biophysics laboratories.

Calculating the temperature dependence of viscosity

For pure water, a 4-term polynomial function recommended by The International Association for the Properties of Water and Steam (ref. 76) is used to calculate the viscosity as a function of temperature:

Equation 30:

$$\eta_T = 10^{-5} \times \left(280.68 \times t^{-1.9} + 511.45 \times t^{-7.7} + 61.131 \times t^{-19.6} + 0.45903 \times t^{-40}\right)$$

where η_T is the viscosity (in centipoise) at temperature T (in Kelvin), and $t = T/300$. This formula is accurate within the uncertainty of the underlying data over the entire liquid range.

The viscosity of D_2O *vs.* temperature is calculated using a formula from Cho *et al.* 1999 (ref. 82):

Equation 31:

$$\eta_{D_2O,T} = 8.8560402\left[\Delta T + 2.7990 \times 10^{-3}(\Delta T)^2 - 1.6342 \times 10^{-5}(\Delta T)^3 + 2.9068 \times 10^{-8}(\Delta T)^4\right]^{1.55255}$$

where ΔT = (T - 231.832) for T in Kelvin. This formula is accurate to better than 0.1% over the entire liquid range32

For H_2O^{18} and D_2O^{18} the experimental data of Wolf & Kudish (ref 96) unfortunately only cover the temperature range from 15 to 35 ^oC. From those experimental data the ratios of the viscosity for the O^{18} isotope to the corresponding O^{16} isotope was calculated at each temperature, and then each set of ratios was fitted as a linear function. For H_2O^{18} that linear fit provided a good fit of the data, with deviations consistent with the estimated experimental uncertainty of 0.0006 cp. Therefore for H_2O^{18} the viscosity as a function of temperature is calculated using those linear fit parameters by

Equation 32:

$$\eta_{H_2O^{18},T} = \eta_{\text{water},T} \times \left[1.0535 - 0.000111(T - 298.15)\right]$$

where the viscosity of water is calculated as noted above (ref. 76).

However for D_2O^{18} the experimental viscosity ratios exhibit higher variability, and are not well fitted as a linear function, with the slope from such a fit being statistically not significantly different from zero. Therefore the viscosity ratio relative to D_2O was simply fixed at the mean value from the 5 experimental temperatures, 1.0462. That is, for D_2O^{18} the viscosity as a function of temperature is calculated by

Equation 33:

$$\eta_{D_2O^{18}}(T) = \eta_{D_2O}(T) \times 1.0462$$

where the viscosity of D_2O is calculated as noted above (ref. 82).

The viscosity for mixtures of different isotopic forms of water are calculated assuming there is no volume change upon mixing (Ref. 45), which means the viscosity for each isotopic form is weighted according to its volume fraction.

CAUTION: Although SEDNTERP will extrapolate viscosities for buffers containing H_2O^{18} or D_2O^{18} to temperatures outside the range from 15 to 35 ^oC where experimental data exist, the accuracy of such extrapolated values is suspect at best.

Estimating a buffer viscosity from its composition

To interpolate a viscosity for a buffer press the **From Composition button** within

the Viscosity panel on the Buffer Properties major tab. The Buffer Composition form will then appear. This same form will also allow you to calculate the buffer density and refractive index, and optionally use those calculated values as the properties of your buffer.

The first step SEDNTERP uses to calculate the viscosity of a buffer is to calculate the viscosity of the buffer at 20 ^{O}C, using "viscosity increments". Then, the viscosity is corrected for temperature assuming that water is the predominant component in the buffer. While that assumption is true for solutions containing moderate amounts of other components, significant errors may be introduced for solutions containing high solute concentrations. For such solutions at temperatures far from 20 ^{O}C, the user is advised to measure the viscosity and enter it manually.

Calculating viscosity increments

Tables of viscosities of various single solute solutions at various concentrations were fitted by least squares methods to a common polynomial form. The sets of coefficients for each solute are stored in the program database. From this, the viscosity of a single solute solution anywhere within the original fit concentration range can be interpolated. The viscosity increment of any concentration of buffer component is simply:

Equation 34:

$$\Delta\eta_i = 1.0016 \times \left[a_i + b_i \cdot 10^{-3} C^{1/2} + c_i \cdot 10^{-2} C + d_i \cdot 10^{-3} C^2 + e_i \cdot 10^{-4} C^3 - 1.00 \right]$$

where $\Delta\eta_i$ is the viscosity increment of the i^{th} component with molar concentration C, a-e are the coefficients of the least squares fit of the original viscosity data. The multiplicative factor of 1.0016 (the viscosity of water at 20 ^{O}C in cp) is needed because the original fits were made from data based on the relative viscosity of water = 1.00.

The inclusion of the square root term improves the fit at low concentrations for some solutes. In principle, the value of a should be exactly 1, but substantially improved fits are sometimes obtained when this term is allowed to vary. Thus, equation 34 is used as an empirical fitting function, and, as such, no physical interpretation of the parameters should be made. For each component, the parameters from the most parsimonious polynomial providing an adequate fit are used by SEDNTERP. Adequacy of the fits were judged by the absence of systematic patterns in plots of the residuals *vs.* concentration.

The concentration range over which each set of coefficients is valid, and the maximum error of the fit over that concentration range, are recorded in the program database. The precision of the viscosity data in the tables is generally much lower than that for the density data, and most solutes show maximum errors of 0.05 to 0.1%, with a few reaching up to ~0.7%.

Having computed the viscosity increments for each component, SEDNTERP will calculate the

viscosity of the buffer at 20 ^{O}C (including the effects of any heavy isotopes of water) by simply adding the viscosity increments to the viscosity of the current isotopic mixture of water, η_{heavy}, at 20 ^{O}C:

Equation 35:

$$\eta_{buffer,\,20} = \eta_{heavy,\,20} + \sum_i \Delta\eta_{C_i}$$

Finally, the buffer viscosity at the desired temperature *T* is calculated using:

Equation 36:

$$\eta_{buffer,\,T} = \eta_{buffer,\,20} \times \frac{\eta_{heavy,\,T}}{\eta_{heavy,\,20}}$$

where both $\eta_{heavy,T}$ and $\eta_{heavy,20}$ are calculated to include the effects of any heavy isotopes of water in this buffer.

3.3.6 Interpolating a Buffer Refractive Index

The refractive index, *n*, of the buffer will have a direct and major influence on light scattering studies (including diffusion coefficients measured by dynamic light scattering). It also indirectly influences AUC experiments using the Rayleigh interferometer, since differences in buffer refractive index will change the macromolecule's specific refractive increment *dn/dc*, which in turn changes the conversion from signal units (fringes) to weight concentration units (the concentration calibration).

CAUTION: Be aware that SEDNTERP calculates and reports the refractive index **relative to vacuum**, not relative to air. At 20 °C and 589.26 nm the refractive index of dry air is 0.000272.

Calculating the refractive index of water at a particular temperature and wavelength

An equation from Harvey et al (1998) (ref. 83) is used to calculate the refractive index of pure water (relative to vacuum) as a function of temperature and wavelength. This equation uses the density of water $\rho(T)$ (calculated as described here), reference constants T^* = 273.15 K, ρ^* = 1 g/mL, and λ^*

= 589 nm, and dimensionless variables $\overline{T} = T/T^*;\quad \overline{\rho} = \rho/\rho^*;\quad \overline{\lambda} = \lambda/\lambda^*$.

Equation 37:

$$\frac{n^2-1}{(n^2+2)\overline{\rho}} = a_0 + a_1\overline{\rho} + a_2\overline{T} + a_3\overline{\lambda}^2\overline{T} + a_4\overline{\lambda}^{-2} + \frac{a_5}{\overline{\lambda}^2 - \overline{\lambda}_{UV}^{\,2}} + \frac{a_6}{\overline{\lambda}^2 - \overline{\lambda}_{IR}^{\,2}} + a_7\overline{\rho}^2$$

The values for coefficients a_0 to a_7 and wavelength parameters $\overline{\lambda}_{UV}$ and $\overline{\lambda}_{IR}$ are listed in the table below. This formula for water is accurate to better than ± 0.00006 over the entire liquid range and from 400 to 700 nm.

Coefficients and parameters for Equation 37:

$a_0 = 2.44257733 \times 10^{-1}$	$a_4 = 1.58920570 \times 10^{-3}$
$a_1 = 9.7634476 \times 10^{-3}$	$a_5 = 2.45934259 \times 10^{-3}$
$a_2 = -3.73234996 \times 10^{-3}$	$a_6 = 9.00704920 \times 10^{-1}$
$a_3 = 2.68678472 \times 10^{-4}$	$a_7 = -1.66626219 \times 10^{-2}$
$\overline{\lambda}_{UV} = 2.292020 \times 10^{-1}$	$\overline{\lambda}_{IR} = 5.432937$

Calculating the refractive index of D_2O

For calculating the refractive index of D_2O as a function of temperature and wavelength a formula from Odhner & Jacobs (2012) (ref. 84), is used:

Equation 38:

$$n_{\lambda,T} = 1.0244 + 3329.2/\lambda^2 + 2.6048\times10^{-3}\cdot T - 1.630\cdot T/\lambda^2 - 7.248\times10^{-6}\cdot T^2 + 6.15\times10^{-9}\cdot T^3$$

where *T* is in Kelvin and λ is in nm. This formula is accurate within ± 0.0002 over the visible range and from 5 to 85 °C. Note, however, that this formula returns the refractive index of D_2O *relative to air*. Therefore, the small correction for the RI of air is computed using this formula for the refractive index of dry air at 1 atm and 15 °C from Ciddor (1996):

Equation 39:

$$10^8\left(n_{\text{air},15}-1\right)=\frac{2.380185\times10^8}{\left(5.792105-\frac{1}{\lambda^2}\right)}+\frac{1.67917\times10^{11}}{\left(5.7362\times10^{-4}-\frac{1}{\lambda^2}\right)}$$

This computed value is then adjusted for temperature assuming tracks the change in the density of air from the ideal gas law, and that final value for air is then added to the value from Equation 38 above to get the final RI value for pure D_2O.

No similar data for the RI of H_2O^{18} or D_2O^{18} have been located, so the RI of buffers containing those isotopes cannot currently be calculated. However, given that the RI of D_20 and H_2O differ by only 0.36% at 589 nm, for many purposes it may be reasonable to assume that the RI for H_2O^{18} or D_2O^{18} is negligibly different from that for ordinary water.

Estimating the refractive index from the buffer composition

To interpolate a refractive index for a buffer press the **From Composition button** within the Refractive index panel on the Buffer Properties major tab. The Buffer Composition form will then appear. This same form will also allow you to calculate the buffer density and viscosity, and optionally use those calculated property values as the properties of your buffer.

The first step SEDNTERP uses to calculate the refractive index (RI) of a buffer is to calculate the refractive index of the buffer at 25 ^oC and a wavelength of 589.26 nm using 'refractive index increments'. Then, the refractive index is corrected for temperature and/or wavelength assuming that water is the predominant component in the buffer. While that assumption is true for solutions containing moderate amounts of other components, significant errors may be introduced for solutions containing high solute concentrations (Ref. 8).

Calculating refractive index increments

Tables of the refractive index (at 20 ^oC and 589.26 nm) of single solute solutions at various concentrations were located for 59 of the 102 compounds currently in the program database. Those tables were then fitted by least squares methods to a common polynomial form (essentially the same polynomial used for determining density increments). The sets of coefficients for each

solute are stored in the program database. From this, the refractive index of a single solute solution anywhere within the original fit concentration range can be interpolated. The refractive index increment of any concentration of buffer component is simply:

Equation 40:

$$\Delta n = a + b \cdot 10^{-3} C^{1/2} + c \cdot 10^{-2} C + d \cdot 10^{-3} C^2 + e \cdot 10^{-4} C^3 + f \cdot 10^{-6} C^4 - n_{water}$$

where Δn is the refractive index increment at a certain concentration, *a*-*f* are the coefficients from the polynomial fit using least squares methods (Ref. 46), *C* is the molar concentration of solute, and n_{water} is the refractive index of pure water.

The inclusion of the square root term improves the fit at low concentrations for some solutes, but for the majority this coefficient was fixed at zero. In principle, *a* should be the refractive index of pure water at 20 ^{o}C, but substantially improved fits are sometimes obtained when this term is allowed to vary. Thus, equation 40 is used as an empirical fitting function, and, as such, no physical interpretation of the parameters should be made. For each component, the parameters from the most parsimonious polynomial providing an adequate fit are used by SEDNTERP. Adequacy of the fits were judged by the following criteria: 1) the r.m.s. error of the residuals should correspond to the last decimal place of the tabulated values; and 2) graphs of the residuals as a function of either the dependent or independent variable should exhibit no systematic deviations with magnitudes greater than the fourth decimal place.

The concentration range over which each set of coefficients is valid, and the maximum error of the fit over that concentration range, are recorded in the program database. Only one solute among those included when the program is installed has a maximum error exceeding 0.0001 (trifluoracetic acid), and even for TFA that maximum error is only 0.00022.

Having computed the refractive index increments for each component, SEDNTERP will calculate the refractive index of the buffer at 20 ^{o}C by simple addition of the refractive index increments:

Equation 41:

$$n_{\text{buffer, 20, 589.26}} = n_{\text{heavy, 20, 589.26}} + \sum_i \Delta n_{C_i}$$

Finally, the buffer RI at the desired temperature *T* and wavelength *λ* is calculated using:

Equation 42:

$$n_{\text{buffer}, T, \lambda} = n_{\text{buffer, 20, 589.26}} \times \frac{n_{\text{heavy}, T, \lambda}}{n_{\text{heavy, 20, 589.26}}}$$

where both $n_{heavy,\ 20,\ 589.26}$ and $n_{heavy,\ T,\ \lambda}$ include the effects of any D_2O in this buffer.

3.3.7 Calculating an Equivalent Radius

The starting point for hydrodynamic analysis from sedimentation or diffusion data is the calculation of the expected radius for a rigid, impermeable and incompressible sphere having the identical volume as the protein of interest. This quantity is called the "equivalent radius" and is designated R_0. As will be explained below, there is some ambiguity as to exactly how R_0 should be determined.

Don't confuse the equivalent radius with the 'Stokes' or 'hydrodynamic' radius, a property which does reflect both the actual shape of the molecule and its hydration.

SEDNTERP automatically calculates the equivalent radius once a value for the molar mass is available. SEDNTERP also then calculates the minimal frictional coefficient from the equivalent radius, and from that minimal frictional coefficient the values of S_{max} (the maximum possible sedimentation coefficient) and D_{max} (the maximum possible diffusion coefficient) are also calculated (see Calculating Smax and Dmax).

The equivalent radius reported on the Derived Hydrodynamic Results major tab is either that from the "Teller method" or the "vbar method" (both discussed in detail below), depending on which of two radio buttons is selected.

Following the calculation of the equivalent radius, other estimates of physical properties of the sample can be calculated. These are discussed generally in Calculating Estimates of Hydration and Asymmetry, in Calculating the Minimum Frictional Coefficient, and more particularly in the Help sections dealing with the individual boxes on the Derived Hydrodynamic Results major tab.

The "vbar" method

In order to account for the fact that the protein is displacing solvent rather than occupying a vacuum, the value of R_0 often is calculated via

Equation 46:

$$R_0 = \sqrt[3]{\frac{3M\bar{v}}{4\pi N_0}}$$

which gives R_0 in units of cm when M is in Daltons and $\bar{v}$ is in mL/g.

However, there is no theoretical justification for the use of $\bar{v}$ in this calculation (Ref. 9) except that $\bar{v}$ reflects, in large part, the volume of solvent displaced per gram of the protein. It is often found that R_0 calculated using equation 41 is too small, thus leading to overestimates of the asymmetry (Refs. 30, 33). This is not surprising since $\bar{v}$ includes solute-solvent interactions other than solvent displacement. Significant errors can be introduced if, for example, there is considerable electrostriction of the solvent by a highly charged protein (Refs. 6, 34). Moreover, this calculation under-estimates the contribution surface roughness (rugosity) makes to the protein hydrodynamics (Ref. 30).

The Teller method

Since there is no theoretical justification for equation 41, it has been suggested that R_0 can instead be calculated using an empirical relationship based on protein volumes observed in X-ray crystallographic structures (Refs. 28, 33). Teller has described two such apparent radii (Ref. 33). The first, R_p, can be calculated by

Equation 47:

$$R_p = 6.72 \times 10^{-9} \sqrt[3]{M}$$

which gives R_p in units of cm when M is in Daltons.

R_p is determined from the packing volume in protein-protein complexes (and hence we assign the subscript "p" to this radius to distinguish it from R_0 calculated with Eq. 41). This formula implies an effective "packing" $\bar{v}$ (which we will give the symbol $\bar{v}_p$) of 0.7654 mL/g.

Teller also proposed a second estimate of an effective radius, R_s, that accounts for the surface roughness of the protein, which is given (again in units of cm) by

Equation 48:

$$R_s = 9.405 \times 10^{-9} \sqrt[3]{M}$$

Thus, R_s = 1.4 x R_p . However, the use of R_s in hydrodynamic calculations leads to absurdly low estimates of the hydration and axial ratio (unpublished data). Other authors have used R_p as an estimate of the anhydrous radius, but have generally neglected R_s, apparently for similar reasons (Ref. 28). Use of these relationships typically yield larger values than R_0 and, consequently, smaller estimates of the axial ratios.

Regardless of whether equation 41 or 42 is used, R_0 and R_p are considered to be the minimum radii possible for the anhydrous protein (Refs. 28, 29, 30, 31, 32, 33). All evidence to date suggest that R_p provides better estimates of the anhydrous radius.

3.3.8 Calculating the Minimum Frictional Coefficient

The predicted minimum frictional coefficient for a molecule at 20 °C (normally given the symbol f_0) is calculated from the Stokes-Einstein relationship using the calculated equivalent radius (either R_0 or R_p) in

Equation 43:

$$f_0 = 6\pi\eta_{20,w}R_0 \text{ or } f_p = 6\pi\eta_{20,w}R_p$$

where $\eta_{20,w}$ is the standard viscosity of pure water at 20 °C.

NOTE: For notational purposes, SEDNTERP labels the minimum frictional coefficient as f_p when it has been computed from R_p (*i.e.* via the Teller method). However, when reported in a publication it would usually be given the symbol f_0.

SEDNTERP then uses the minimum frictional coefficient to calculate both a maximum sedimentation coefficient (S_{max}) and a maximum diffusion coefficient (D_{max}). It is also used in Calculating the Perrin P function and the f(hyd)/f0 Ratio, which in turn is used in calculating the degree of molecular asymmetry for ellipsoid or cylinder models.

3.3.9 Calculating Smax and Dmax

For any molecule, the maximum possible value for its sedimentation and diffusion coefficients in

water at 20 °C, S_{max} and D_{max}. will arise when that molecule is a sphere and has zero hydration (*i.e.* an anhydrous sphere with the equivalent radius). That condition corresponds to the minimum frictional coefficient, f_0.

When a value for f_0 is available we can calculate S_{max} using

Equation 44:

$$S_{max} = \frac{M\left(1-\bar{v}_{20}\rho_{20,w}\right)}{N_0 f_0} = \frac{M_b}{N_0 f_0}$$

where $\bar{v}_{20}$ is the partial specific volume at 20 °C, and $\rho_{20,w}$ is the density of pure water at 20 °C = 0.998213 g/mL (ref. 76).

D_{max} can be calculated using:

Equation 45:

$$D_{max} = \frac{RT}{N_0 f_0}$$

where T in this case is 293.15 °K.

When the equivalent radius is being calculated using the Teller method then f_p is substituted for f_0 in Equations 44 and 45.

3.3.10 Converting s* to s0(20,w)

The sedimentation coefficient provides a rigorous hydrodynamic descriptor of a molecule. However, to be most useful in this regard, and prior to further interpretation, *s** must be standardized and corrected for concentration effects. By convention, after *s** has been scaled to standardize it to conditions corresponding to pure water at 20 ^{o}C, and extrapolated to zero protein concentration, the resulting value is designated $\boldsymbol{s^0_{20,w}}$. Because of the solvent-independent nature of $s^0_{20,w}$, it describes quantitatively the fundamental hydrodynamic properties of the protein, and it is this value which is most useful in comparing the sedimentation behavior of

different proteins. Moreover SEDNTERP must use $s^0_{20,w}$ in calculating the limits of hydration and when estimating the asymmetry of a macromolecule. Finally, comparisons of $s^0_{20,w}$ determined for a molecule in solvents of differing composition can yield unique information concerning changes in macromolecular interactions, shapes, sizes and hydration.

Standardization of raw experimental sedimentation coefficients involves corrections to account for the buffer density ρ and its viscosity η, as well as the protein's partial specific volume $\bar{v}$. Such corrections are needed because ρ and $\bar{v}$ affect s^* by their presence in the buoyancy correction term $(1-\bar{v}\rho)$ (Eqn. 1), while η affects s^* inversely through the frictional coefficient f.

Standardizing s* is usually a two-step process:

1. Correct s* to the value that would theoretically be observed in pure water at 20 ^{o}C, yielding a value that is by convention designated as $\boldsymbol{s_{20,w}}$. That is, this step corrects for the differences between the actual buffer density and viscosity versus the density and viscosity of water, as well as for any temperature difference away from 20 ^{o}C.
2. Measure s^* at several protein concentration, standardize those experimental values to $s_{20,w}$ values, and then extrapolate to zero protein concentration (yielding the final $s^0_{20,w}$ value). Note however that this step may not be required---see the 'Tip' box below.

1. Standardizing s*:

s^* is corrected to standard conditions of water at 20 ^{o}C using:

Equation 11:

$$s_{20,w} = s^* \frac{\left(1-\overline{v}_{20}\rho_{20,w}\right)}{\left(1-\overline{v}_T\rho_{T,b}\right)_{T,b}} \frac{\eta_{T,b}}{\eta_{20,w}}$$

where the subscripts refer to the experimental temperature (T, 20=20 ^{o}C) and buffer conditions (b=buffer, w=water), respectively. Note that the $\bar{v}$ in the denominator should be corrected for specific ion effects whenever possible (Ref. 6). The current "best" values for water at 20 C are $\rho_{20,w}$ = 0.998213 g/mL (ref. 76) and $\eta_{20,w}$ = 1.0016 cp (ref 77).

2. Accounting for the concentration-dependence of s*:

Prior to further analysis, $s_{20,w}$ should be determined for a number of samples at different initial protein concentrations. A graph then should be constructed to extrapolate $s_{20,w}$ to zero protein concentration. The nomenclature for a standardized, concentration-corrected (i.e. lim $c \rightarrow 0$) sedimentation coefficient is $s^0_{20,w}$. In cases where well-resolved boundaries are being examined, such corrections should be made for each component individually, if at all possible.

The determination of $s_{20,w}$ at various solute concentrations may reveal either increases or decreases in s with increasing solute concentration. A decrease in $s_{20,w}$ with increasing c is expected on hydrodynamic grounds (Refs. 21, 22). A rather weak dependence (about a 1% change in s per mg/mL) is observed for spherical molecules, whereas asymmetric molecules will exhibit a much stronger concentration dependence (Ref. 23).

Increases in $s_{20,w}$ with increasing solute concentration, or the appearance and growth of faster-sedimenting boundaries at increased solute concentrations, is definitive proof for mass-action association. The full analysis of sedimentation velocity data from interacting systems is reviewed in "Analytical Ultracentrifugation in Biochemistry and Polymer Science" and elsewhere (Refs. 17, 24).

In many cases, the effect of protein concentration on $s_{20,w}$ can be taken into account using equation 5:

$$s^0_{20,w} = s_{20,w} / (1 - k_s \times c)$$

where the protein concentration c is in mg/mL and $\boldsymbol{k_s}$ is an empirical coefficient that characterizes the strength of the concentration dependence. Ideally the coefficient k_s should be measured by fitting $s_{20,w}$ values measured at several concentrations to the equation above. However it is reported that k_s is equal to 9 mL/g for spherical proteins (refs. 21, 22, 27, 28), and that is the default value assumed by SEDNTERP.

SEDNTERP will automatically perform this calculation (applying Eq. 5) if the user enters the concentration of protein in mg/mL into the text box on the Sedimentation Velocity minor tab on the main form. The user may also evaluate k_s by fitting experimental sedimentation coefficients at 3 or more concentrations to Eq. 5 by clicking on the **Evaluate k_s button** located on either the Sample main tab, or the Sedimentation Velocity minor tab of the Main form. Both of those buttons then bring up the Evaluate ks and Extrapolate s0(20,w) form.

Alternatively the user can input a known experimental (or other) value for k_s by clicking on the **Use Known Value button** located on the Sample main tab, or by directly entering a k_s value into a text box located on the Sedimentation Velocity minor tab of the Main form.

Note that much larger values of k_s are expected for asymmetric proteins, and therefore the default k_s value of 9 mL/g should be used only if it is known that the protein is spherical or nearly

so (or if the protein concentration is so low that the correction is comparable to the likely uncertainty in the $s_{20,w}$ value). Note that much larger values of k_s are also expected for highly-charged molecules at low ionic strength. There has not been a systematic study of the variation in k_s with axial ratio, and this might make a useful area for study.

TIP: Modern AUC instruments are often able to measure the sedimentation coefficient at a concentration low enough that the concentration dependence of s* is negligible compared to the precision of the s* values (or compared to the precision needed for the experimenter's purpose). For example, we can normally easily do velocity experiments at protein concentrations of 0.5 mg/mL or less, which for a globular protein means that the difference in s* relative to infinite dilution is only roughly ~0.5%, and therefore it *may* be reasonable to omit the extrapolation. On the other hand, if the protein is mostly unfolded or highly asymmetric in shape, or if it has a relatively high electric charge and the buffer ionic strength is very low, the concentration dependence could be an order of magnitude larger (or more), and thus we would need to go to an order of magnitude lower concentration to make the concentration dependence negligible. Similarly, for nucleic acids the concentration dependence is quite large even in buffers of physiological ionic strength.

Remember though that it is always better to not make assumptions, but rather directly measure the concentration dependence, and then make the extrapolation to zero concentration.

3.3.11 Converting D* to D0(20,w)

The diffusion coefficient provides a rigorous hydrodynamic descriptor of a molecule. However, to be most useful in this regard, and prior to further interpretation, D^* must be standardized and corrected for concentration effects. By convention, after D^* has been scaled to standardize it to conditions corresponding to pure water at 20 ^{o}C, and extrapolated to zero protein concentration, the resulting value is designated $\boldsymbol{D^0_{20,w}}$. Because of the solvent-independent nature of $D^0_{20,w}$, it describes quantitatively the fundamental hydrodynamic properties of the protein, and it is this value which is most useful in comparing the diffusion behavior of different proteins. Moreover SEDNTERP must use $D^0_{20,w}$ in calculating the limits of hydration or when estimating the asymmetry of a macromolecule. Finally, comparisons of $D^0_{20,w}$ determined for a molecule in

solvents of differing composition can yield unique information concerning changes in macromolecular interactions, shapes, sizes and hydration.

Standardization of raw experimental diffusion coefficients involves corrections to account for the solvent viscosity η as well as the experiment temperature T. Such corrections are needed because η affects D^* inversely through the frictional coefficient f, while T (in Kelvin) affects D^* linearly (because increased thermal energy increases the diffusion rate).

Standardizing D* is usually a two-step process:

1. Correct D* to the value that would theoretically be observed in pure water at 20 ^{O}C, yielding a value that is by convention designated as $\boldsymbol{D_{20,w}}$. That is, this step corrects for the differences between the actual buffer viscosity versus the viscosity of water, as well as for any temperature difference away from 20 ^{O}C.
2. Measure D^* at several protein concentration, standardize those experimental values to $D_{20,w}$ values, and then extrapolate to zero protein concentration (yielding the final $D^0_{20,w}$ value).

1. Standardizing D*:

D^* is corrected to standard conditions of water at 20 ^{O}C using:

Equation 13:

$$D_{20,w} = D^* \frac{\eta_{T,b}}{\eta_{20,w}} \frac{293.15}{T}$$

where the subscripts refer to the experimental temperature (T, 20 ^{O}C) and buffer conditions (b=buffer, w=water), respectively. The current "best" viscosity value for water at 20 C is $\eta_{20,w}$ = 1.0016 cp (ref 77).

2. Accounting for the concentration-dependence of D*:

Prior to further analysis, $D_{20,w}$ should be determined for a number of samples at different initial protein concentrations. A graph then should be constructed to extrapolate $D_{20,w}$ to zero protein concentration. The nomenclature for a standardized, concentration-corrected (i.e. lim $c \rightarrow 0$) diffusion coefficient is $D^0_{20,w}$. In cases where well-resolved boundaries are being examined, such corrections should be made for each component individually, if at all possible.

The determination of $D_{20,w}$ at various solute concentrations may reveal either increases or decreases in D with increasing solute concentration. An increase in $D_{20,w}$ with increasing c is

expected on thermodynamic grounds (Refs. 73, 74).

The concentration dependence of the diffusion coefficient is usually characterized by an empirical parameter $\boldsymbol{k_D}$, which is essentially defined by the relation:

Equation 6:

$$D(c) = D_0\left(1 + k_D \cdot c\right)$$

where $D(c)$ is the diffusion coefficient (either raw or standardized) at finite weight concentration c, and D_0 the value that would be observed at infinite dilution.

SEDNTERP will automatically perform this calculation (applying Eq. 6) if the user enters the concentration of protein in mg/mL into the text box on the Diffusion minor tab on the main form. The user may also evaluate k_D by fitting experimental diffusion coefficients at 3 or more concentrations to Eq. 6 by clicking on the **Evaluate k_D button** located on either the Sample main tab, the Diffusion minor tab, or the Sedimentation Velocity minor tab of the Main form. Both of those buttons then bring up the Evaluate kD and Extrapolate D0(20,w) form.

Alternatively the user can input a known experimental (or other) value for k_D by clicking on the **Use Known Value button** located on the Sample main tab, or by directly entering a k_D value into a text box located on the Diffusion or Sedimentation Velocity minor tabs of the Main form. (Note that the buttons and text box related to k_D only appear on the Sedimentation Velocity minor tab when the user indicates that diffusion coefficients were determined from the velocity experiment.)

As discussed in the Concentration Dependence of the Diffusion Coefficient topic, reported values for k_D vary over a wide range (both positive and negative), and there is significant disagreement among theoretical approaches to calculating or predicting k_D. There is also both theoretical and experimental data to suggest that k_D values may differ substantially depending on whether the experimental method involves diffusion driven by a concentration gradient versus without a concentration gradient (as in dynamic light scattering or single particle tracking measurements).

3.3.12 Calculating the Standardized Experimental Frictional Coefficient, f

The translational frictional coefficient f can be calculated from either the sedimentation coefficient or the diffusion coefficient, using Equations 3 or 7, respectively:

Equation 3:

$$s = \frac{M(1-\bar{v}\rho)}{N_0 f} = \frac{M_b}{N_0 f}$$

Equation 7:

$$D^* = \frac{RT}{N_o f} = \frac{k_B T}{f}$$

However that resultant value of *f* depends on the viscosity of the buffer used for the experiment (it is proportional to the viscosity). Therefore before trying to interpret *f* (for example comparing values between molecules) it is desirable to correct the value to a standard condition, where the viscosity is that for water at 20 °C (in exact analogy with what is done to standardize sedimentation and diffusion coefficients). Thus the standardized *f* can be calculated by simply using the $s^0_{20,w}$ or $D^0_{20,w}$ value in Eqs. 3 or 7, respectively.

Such a standardized value for *f* is what is displayed on the Derived Hydrodynamic Results major tab. It could (and probably should) be labeled as an "$f_{20,w}$" value, however that is not the common practice in this field. The value for f_0 or f_p displayed on that page is also standardized.

Note that standardization of frictional coefficients is not needed when they are used to calculate ratios such as f/f_0 or f_{hyd}/f_0 , as long as both values in the ratio refer to the same buffer conditions.

3.3.13 Calculating Maximum Hydration

The predicted *maximum possible* degree of hydration (g-H_2O per g-protein) can be calculated by assuming the protein molecule is a perfect sphere, and that all the excess friction (the difference between *f* and f_o or f_p) is due to hydration.

When the equivalent radius R_0 is computed via the $\bar{v}$ method, then the maximum hydration $\delta_{1,max}$ is calculated by

Equation 51:

$$\delta_{1,\max} = \left[\left(f/f_0\right)^3 - 1\right] \cdot \bar{v} \cdot \rho_{hyd}$$

[note that this equation is incorrect in ref. 59]

where f is the experimental frictional coefficient (see Calculating the Standardized Experimental Frictional Coefficient), f_0 is the minimum frictional coefficient (see Calculating the Minimum Frictional Coefficient), $\bar{\nu}$ is the partial specific volume of the macromolecule, and ρ_{hyd} is the density of the bound hydration water at 20 °C (which is not necessarily the same as that of the bulk water, but is typically assumed to be so).

Alternatively, when R_0 is calculated via the Teller method, $\delta_{1,max}$ can be calculated by substituting f_p for f_0 in equation 51, and setting $\bar{\nu}$ at the 'packing' value of 0.7654 mL/g.

SEDNTERP performs either variation of this calculation automatically on the Derived Hydrodynamic Results major tab when the necessary information has been entered (unless the user has chosen to calculate or enter an estimate of hydration in order to compute a more accurate *a/b* ratio).

Note: The assumed density of the bound hydration water ρ_{hyd} (at the standard temperature of 20 °C) is a program default value. When the program is installed that value is the same as that for pure bulk water, but the user may alter that value using the **Change Program Defaults...** command on the File Menu.

3.3.14 Estimating Hydration

Many protein scientists handle hydration by simply assuming a 'typical' or 'default' value for the hydration, δ_1, often in the range from 0.25 to 0.4 g/g. SEDNTERP allows a 'default' state for δ_1 (with an initial value of 0.3 g/g), and it can use that default in other hydrodynamic calculations. However, there is limited theoretical or empirical justification for this practice.

Water of hydration is commonly defined to be water that differs from the bulk phase due to its interaction with the protein. Our knowledge of the degree of hydration for proteins is quite limited, largely because the reported values for δ_1 are dependent on the technique used to measure it (Refs. 28, 29, 30, 31, 32, 33, 39, 40, 41, 42, 43, 51, 52, 53, 54). This variability arises because different techniques measure different amounts of water "under the influence" of the protein since they are detecting fundamentally different properties of that water.

Most methods for measuring protein hydration distinguish two different realms of hydration water: (1) water directly interacting with the protein (the primary hydration shell), and (2) a secondary shell of water that is "influenced" by being in contact with the primary shell (Ref. 52). It is not clear whether one, both, or some combination of these two shells corresponds to the hydration that influences sedimentation and diffusion behavior.

A significant limitation of this interpretation is that it fails to take into consideration that the

hydration that is relevant for sedimentation and diffusion measurements includes water that is "trapped" or “entrained” by the protein, but which is not necessarily in contact with the molecular surface or the primary hydration shell. In particular, that entrained water includes water trapped in small buried cavities, narrow channels, or crevices (Refs. 94, 102, 103). Recently Fleming et al. (2023, Ref. 102) described a method of calculating the amounts of surface shell and entrained water from a known protein structure. Their results indicate that for small globular proteins the first shell represents most of the total hydration, but proteins larger than ~200 kDa have 0.5 g/g or more of entrained water.
This ambiguity about how to define or measure the amount of hydration that is relevant for sedimentation and diffusion complicates the estimation of the asymmetry due to the common effects hydration and shape have on hydrodynamics (Refs. 28, 29).

Estimating hydration from composition

One estimate of the degree of hydration can be made based on the amino acid composition and using the method of Kuntz (Ref. 51). This method relies on the amount of "unfreezable" water being an accurate reflector of hydration as it relates to hydrodynamics. The calculation is simply:

Equation 52:

$$\delta_1 = \frac{18}{M_r} \sum H_i N_i$$

Where H_i is the hydration assigned to each amino acid in the program database and N_i is the number of residues of type i. There are two sets of values for H_i : one set is used at pH 4.0 and below, and the second set is used at all other pH values (and was actually measured at pH 7.0). Estimates of δ_1 calculated using equation 52 are not otherwise adjusted for the effects of pH, but are reported to be valid to within 10% above pH 6. Values for H_i at pH 4 reflect the lower hydration expected of the carboxyl groups when they are protonated. The quantity $18/M_r$ in Eq. 52 converts the value from the relative molar scale to the relative mass scale.

Use of equation 52 assumes that all of the amino acids are exposed to solvent, which for globular proteins clearly cannot be the case, and thus that argument implies that δ_1 would tend to be overestimated. On the other hand, this calculation is only for the amount of water directly bound to the protein, and assumes no water is entrained, which would tend to underestimate δ_1.

Although there have been studies correlating predicted hydration using equation 52 with "non-freezable" water (Refs. 51, 54), there has been minimal systematic exploration of whether these values correlate with hydrodynamic properties. However, empirical evidence suggest that δ_1 calculated with this method is "better than anticipated", and that the values of δ_1 are useful in

determining ϕ_2 from equation 18 (Ref. 42).

Finally, note that most of the carbohydrate conjugates in Table 2 have also been assigned non-zero H_i values. The fairly large hydration of 1.22 g/g for polyethylene glycol (PEG) comes from calorimetry data (Ref. 89) and represents water that is bound to the PEG. Note however that recent studies indicate that PEG entrains a very large amount of water, and that this entrainment increases strongly with the PEG chain length, resulting in an extremely large total hydration of 23.2 g/g for a 40 kDa PEG (Ref. 104).

3.3.15 Calculating Hydration from Asymmetry Estimates

Should the user supply either an ellipsoidal axial ratio (as major-axis/minor-axis), or the length/diameter ratio for a cylinder, the corresponding values of the degree of hydration for prolate or oblate ellipsoid models, or for rod or disk cylinder models, can be calculated by SEDNTERP.

Calculating hydration from known ellipsoidal asymmetry

For ellipsoids, the first step is to calculate a frictional coefficient ratio for an anhydrous ellipsoid having the user-supplied *a/b* major/minor axis ratio, relative to an anhydrous sphere with the same volume, which is the Perrin P function. This ratio will be designated as $P_{prolate}$ for the prolate case and P_{oblate} for the oblate case. These values may be calculated using the formulas of Perrin (Ref. 18). [Note that equations 53-55 below are all incorrect in Ref. 59]:

Equation 53:

$$P_{prolate} = \frac{\left(1-\left[b/a\right]^2\right)^{1/2}}{\left[b/a\right]^{2/3} \cdot \ln\left\{\frac{1+\left(1-\left[b/a\right]^2\right)^{1/2}}{\left[b/a\right]}\right\}}$$

Equation 54:

$$P_{oblate} = \frac{\left(\left[a/b\right]^2 - 1\right)^{1/2}}{\left[a/b\right]^{2/3} \cdot \tan^{-1}\left\{\left(\left[a/b\right]^2 - 1\right)^{1/2}\right\}}$$

The appropriate *P* value is then used to calculate the corresponding degree of hydration using

Equation 55:

$$\delta_1 = \left[\left(\frac{f/f_0}{P}\right)^3 - 1\right] \cdot \bar{v} \cdot \rho_{hyd}$$

where δ_1 is the predicted degree of hydration for the shape model, $\bar{v}$ is the protein's partial specific volume, and ρ_{hyd} is the density of the bound hydration water (which is not necessarily the same as that of the bulk water, but is typically assumed to be so). As when calculating the maximum hydration, f_0 can be calculated using equation 21, or alternatively the empirical equations of Teller (Ref. 33) can be used by substituting R_p for R_0 in computing f_0 and substituting a value of $\bar{v}_p$ = 0.7654 mL/g for $\bar{v}$ (from equation 7 of Ref. 33).

Calculating hydration from known cylindrical asymmetry

For cylinder models, again the first step is to calculate a Perrin P frictional coefficient ratio for an anhydrous rod or disk having the user-supplied *L/d* length/diameter ratio for rods or the *d/H* diameter/height ratio for disks. The value for *P* may be calculated using a formula from Hansen 2004 (ref. 85).

Equation 56:

$P = 1.0304 + 0.0193x + 0.06229x^2 + 0.00476x^3 + 0.00166x^4 + 0.00000266x^7$

where for rods $x = ln(L/d)$ and for disks $x = -ln(d/H)$. This polynomial is stated to be accurate to within 0.1% for *L/d* or *d/H* ratios from 1 to 100.

This *P* value can then simply be substituted into Equation 55 above to calculate δ_1 for rods or disks.

3.3.16 Limitations on Estimating Hydration and Asymmetry

One cannot determine from sedimentation or diffusion experiments alone accurate, unambiguous values for both the degree of hydration (δ_1) and the asymmetry (either the ellipsoid axial ratio *a/b* or the cylinder length/diameter ratio *L/d*) (Refs. 27, 28, 29, 30). However, maximum values for these parameters may be calculated if *s* (or *D*) and *M* are available ("Teller method"). The "vbar method" may also be used for calculating maximum hydration and asymmetry if a value for $\bar{v}$ is also available. In addition, if one of the two parameters (δ_1 or asymmetry ratio) are known independently, estimates of the second parameter can be made. However, at the outset it must be realized that these parameters, especially the *a/b* or *L/d* ratios, are merely descriptors for models and may bear little relevance to the actual molecular structure.
Another limitation in interpreting the effects of hydration on sedimentation and diffusion measurements is uncertainty about the density of the hydration water. Although it has been common practice in this field to assume that hydration water (of any type) has the same density as bulk water, x-ray and neutron scattering data for proteins in solution indicates that the density of the primary hydration layer is 10-15% higher than that of bulk water (Refs. 105, 106). Overall, the uncertainty about how to quantify hydration (discussed here), and the uncertainty about what density to assign to that water of hydration, is often a significant limitation for interpretation of hydrodynamic parameters.

Sources of error

There are two sources of error that must be addressed. First, the calculation of (*a/b* or *L/d*) and δ_1 will rely on an estimate of R_0 or R_p, and, as discussed here, the equations for these quantities are empirical. Second, that radius is then used to calculate f_0 using equation 43. However, equation 43 is derived assuming complete "stick" boundary conditions (i.e. a monolayer of solvent moves with the macromolecule). If complete "slip" boundary conditions are assumed (*i.e.* there is no interaction between the solvent and the macromolecule), equation 23 should be rewritten as equation 70 (Ref. 18).
Equation 70:

$$f_0 = 4\pi\eta_{20,w}R_0 \text{ or } f_p = 4\pi\eta_{20,w}R_p$$

Thus, even though all evidence suggest that a purely "stick" boundary condition prevails, this assumption increases the uncertainty in the f_0 calculation.

3.3.17 Calculating the Perrin P function and the f(hyd)/f0 Ratio

The overall hydrodynamic compactness of a molecule is then often characterized by reporting the ratio f/f_0, where f_0 is the minimum possible frictional coefficient for this molecule (the value for an anhydrous sphere of equivalent volume, see Calculating the Minimum Frictional Coefficient). An f/f_0 ratio of 1.2-1.3 usually indicates a fairly compact structure with low asymmetry and disorder, whereas a ratio above 1.6 suggests an extended structure with fairly high asymmetry (or perhaps a partially-disordered protein).

However, a significant limitation of that ratio is that it fails to distinguish what portion of the ratio is due to hydration, and what portion is due to the shape or flexibility of the molecule. Therefore SEDNTERP now reports the Perrin translational friction function *P* (Ref. 18), which is the ratio of the frictional coefficient for the un-hydrated (anhydrous) molecule to that for an anhydrous sphere with the same volume. **The advantage of the ratio *P* is that it depends *only* on shape or flexibility**.

The name "Perrin translational friction function *P*" is often shortened to "Perrin *P* function" or simply "Perrin *P*". Note too that it is sometimes given the symbol *F* rather than *P*.

SEDNTERP now also the calculates and reports the ratio f_{hyd}/f_0 or f_{hyd}/f_p, which compares the frictional coefficient of a hydrated vs. anhydrous equivalent sphere (and which therefore depends only on hydration). Both *P* and f_{hyd}/f_0 are displayed on the Derived Hydrodynamic Results major tab.

These three different frictional ratios are related through

Equation 49:

$$f/f_0 = P \cdot f_{hyd}/f_0 \text{ ; or } f/f_p = P \cdot f_{hyd}/f_p$$

where f_{hyd}, the frictional coefficient for the hydrated sphere, is calculated by adding a layer of hydration water around the anhydrous equivalent sphere with a specific volume equal to the hydration δ_1 divided by the density of that bound hydration water, ρ_{hyd}. (The density of the bound water is not necessarily the same as that of the bulk water, but is typically assumed to be so).

The value of f_{hyd} is calculated using

Equation 50:

$$f_{hyd} = f_0 \cdot (\bar{v} + \delta_1/\rho_{hyd})^{1/3} \text{ ; or } f_{hyd} = f_p \cdot (\bar{v}_p + \delta_1/\rho_{hyd})^{1/3}$$

Note: The assumed density of the bound hydration water ρ_{hyd} (at the standard temperature of 20 °C) is a program default value. When the program is installed that value is the same as that for pure bulk water, but the user may alter that value using the **Change Program Defaults...** command on the File Menu.

3.3.18 Calculating Ellipsoidal Asymmetry

When no estimate for the hydration δ_1 is available, the *maximum* axial ratio of the prolate ellipsoid of revolution $(a/b)_{p,m}$ and the oblate ellipsoid of revolution $(a/b)_{o,m}$ can be determined using the Perrin *P* function (Ref. 18). In SEDNTERP this is done using a polynomial approximation of the tabulated data for (*a/b*) as a function of the Perrin *P* function from Ref. 18, where the coefficients of the polynomial were determined by least-squares fitting.

This gives the following equations for prolate and oblate ellipsoids, respectively:

Equation 57:

$$(a/b)_{p,m} = 1 + 2.939184 \cdot P^{1/2} + 8.027775 \cdot P + 8.55676 \cdot P^2 - 5.05951 \cdot P^3 + 0.0361508 \cdot P^4$$

Equation 58:

$$(a/b)_{o,m} = 1 + 3.009376 \cdot P^{1/2} + 8.413897 \cdot P + 13.34978 \cdot P^2 - 3.18903 \cdot P^3 + 0.106307 \cdot P^4$$

NOTE: The coefficients in both of the above equations differ from those in Ref. 59; these values give better accuracy.

Over the range 1 <= (a/b) <= 200, these functions are within 0.01 of the true axial ratio, and certainly introduce no error beyond that in estimating *f* or f_0. Note that the maximum axial ratios correspond to assuming the hydration is zero, in which case $P = f/f_0$.

Should an estimate of the degree of hydration be available, the corresponding values for $(a/b)_{p,estimate}$ and $(a/b)_{o,estimate}$ are calculated from equations 57 or 58 using equation 59 below for *P*:

Equation 59:

$$P = \frac{f}{f_0} \cdot \left(\bar{v} + \delta_1 / \rho_{hyd} \right)^{-1/3}$$

where δ_1 is the degree of hydration expressed as g-water/g-protein, $\bar{v}$ is the partial specific volume of the protein, f/f_0 is the ratio of the experimentally determined frictional coefficient to that for the equivalent anhydrous sphere (both at 20 °C), and ρ_{hyd} is the density of the bound hydration water (which is not necessarily the same as that of the bulk water, but is typically assumed to be so). Again, Teller's empirical estimates (Ref. 33) for f_p, and his estimate of $\bar{v}_p$ = 0.7654 mL/g can be substituted for f_0 and $\bar{v}$ in this calculation.

Calculating the ellipsoid dimensions

The final step in the ellipsoid calculations is to calculate the actual length of the major and minor axes. This is done based on the known axial ratio (*a/b*) and by setting the volume of the ellipsoid as the volume of the protein plus the volume of its bound water. The major axis length (in nm) for the prolate and oblate cases, respectively, is given by:

Equation 60:

$$a_p = \left[\frac{3M\left(\bar{v} + \delta_1 / \rho_{hyd} \right)\left(a/b \right)_p^2}{4\pi N_a} \right]^{1/3} \times 10^7$$

Equation 61:

$$a_o = \left[\frac{3M\left(\bar{v} + \delta_1 / \rho_{hyd} \right)\left(a/b \right)_o}{4\pi N_a} \right]^{1/3} \times 10^7$$

Once the major axis length is known, the minor axis length can be calculated from the known axial ratio.

3.3.19 Calculating Cylindrical Asymmetry

The proper hydrodynamic model for molecules that exhibit an elongated shape with a uniform thickness is the rod (a cylinder with length/diameter ratio > 1). For other molecules or complexes the appropriate model may be a disk (a short cylinder with diameter/height ratio > 1).

The Perrin P frictional coefficient ratio for an anhydrous rod or disk with known asymmetry, relative to an anhydrous sphere with the same volume, may be calculated using a formula from Hansen 2004 (ref. 85).

Equation 56:

$P = 1.0304 + 0.0193x + 0.06229x^2 + 0.00476x^3 + 0.00166x^4 + 0.00000266x^7$

where for rods $x = ln(L/d)$ and for disks $x = -ln(d/H)$. This polynomial is stated to be accurate to within 0.1% for L/d or d/H ratios from 1 to 100.

Therefore given an experimental value for P (as calculated and shown on the Derived Hydrodynamic Results major tab) SEDNTERP can solve equation 56 numerically using a Van Wijngaarden-Dekker-Brent root finder (Ref. 101) to find the L/d ratio. When an estimate for δ_1 is available, the estimated rod L/d or disk d/H ratio can be calculated by first determining P from Eq. 59, and then again solving Eq. 56 numerically to find the ratio consistent with that P value.

Calculating the cylinder dimensions

The actual diameter of the cylinder (in nm) can be calculated from the known L/d ratio (substituting $L = H$ for a disk) by setting the total volume of the cylinder as the volume of the protein plus its bound water, and then solving for d, using:

Equation 62:

$$d = \left[\frac{4M\left(\bar{v} + \delta_1/\rho_{hyd}\right)}{\pi N_a \left(L/d\right)} \right]^{1/3} \times 10^7$$

where δ_1 is the degree of hydration expressed as g-water/g-protein, $\bar{v}$ is the partial specific volume of the protein, N_0 is Avogadro's number, M is the molar mass of the macromolecule, and ρ_{hyd} is the density of the bound hydration water (which is not necessarily the same as that of the bulk water, but is typically assumed to be so). Once d is known, L can then be calculated from

the known *L/d* ratio.

If no information about hydration is entered by the user, the degree of hydration is assumed to be zero, and a maximum possible length/diameter ratio, $(L/d)_{max}$, is calculated and reported.

3.3.20 Calculating the Stokes Radius

The Stokes radius is commonly given the symbol R_s. It is also known as the "hydrodynamic radius", and then is often given the symbol R_h. It can be experimentally measured by dynamic light scattering or methods that monitor the Brownian motion of individual particles, and such experimental values should be entered on the Diffusion minor tab. It also can be calculated from either diffusion or sedimentation data, using several different relationships, which are listed below, all of which may be derived from Equation 3 and the Svedberg equation (Eq. 2):

$$R_s = \frac{k_B T}{6\pi\eta D}$$
$$= \frac{M(1-\bar{v}\rho)}{\eta s N_0}$$
$$= \frac{M(1-\bar{v}_{20}\rho_{20,w})R_0}{N_0 f_0 s_{20,w}}$$
$$= \frac{M(1-\bar{v}_{20}\rho_{20,w})R_p}{N_0 f_p s_{20,w}}$$

where the symbols have their usual meaning.

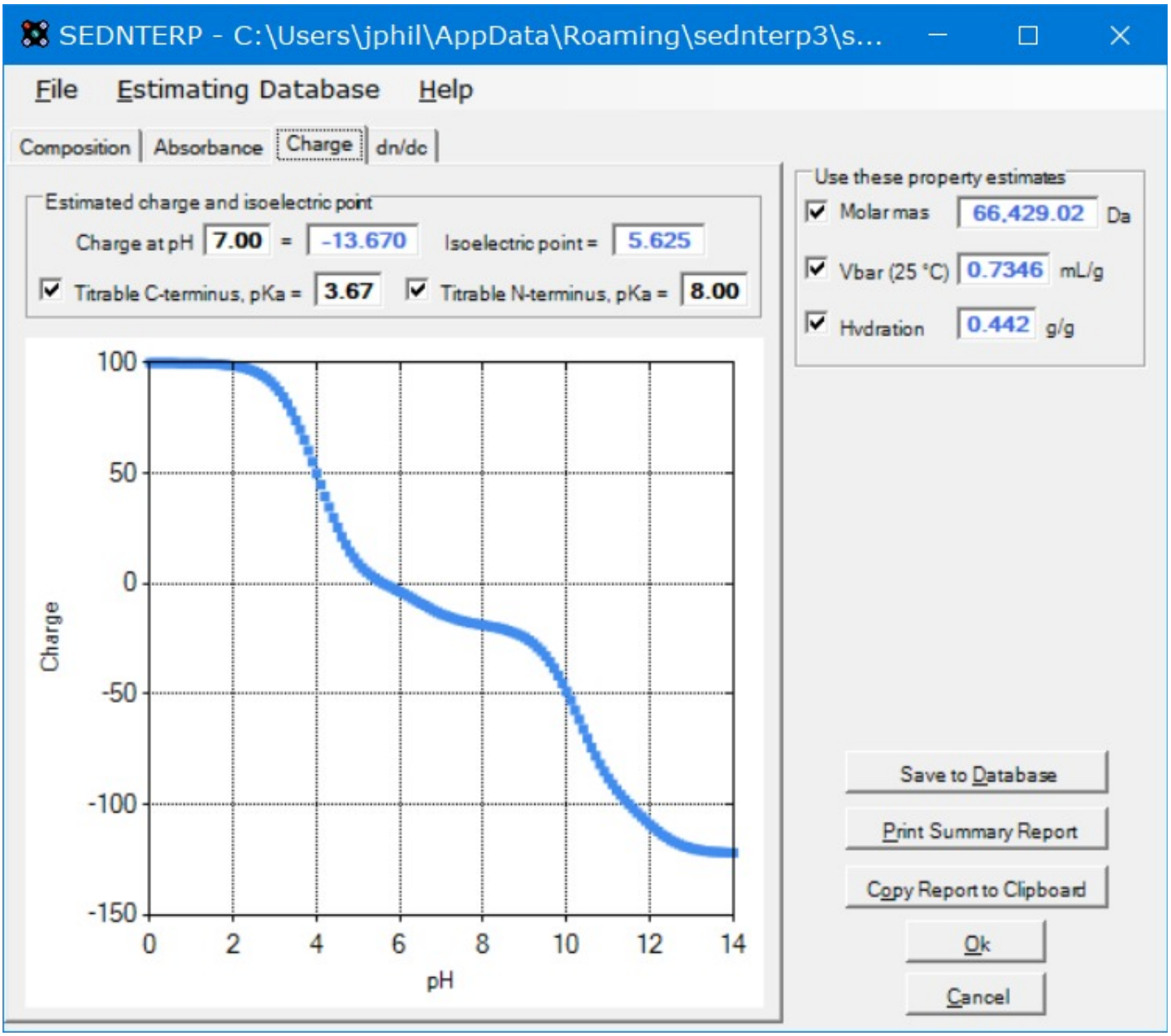
Note that this Stokes radius calculation does *not* depend on whether you choose the vbar or the Teller method for calculating an Equivalent Radius. The calculated Stokes radius is reported on the Derived Hydrodynamic Results major tab.

NOTE: The Stokes radius calculated by SEDNTERP from sedimentation velocity data is, in principle, identical to that measured by dynamic light scattering (DLS) or by single-particle tracking methods that directly image the Brownian motion. However, if the sample has heterogeneity or undergoes reversible association, then different physical methods will likely differ in their sensitivity to the different components, and thus may not yield

the same value for the Stokes radius. In particular, the Stokes radius from DLS may be influenced by trace amounts of un-resolved very large particles which would not be detected in a sedimentation velocity experiment.

3.3.21 Calculating Electric Charge and Isoelectric Point

SEDNTERP will calculate both the estimated total electric charge on a protein (at the current value for the buffer pH) and the estimated isoelectric point (*pI*), based on the composition entered on the Sample Composition form. It will also display a theoretical titration curve (graph of estimated total charge *vs.* pH). These calculations are shown when the Charge tab page of Sample Composition form is displayed, as illustrated below:



Calculating the total charge Z

The estimated charge for each type of ionizable group (whether amino acid or conjugate) at a given pH is calculated using a form of the Henderson-Hasselbalch equation (Refs. 61, Reference 62).

For positively charged groups (*i.e.* when the charge listed in the database is positive) the charge z for the group is given by

Equation 68

$$z = \left(1 + 10^{(pH - pK_a)}\right)^{-1}$$

where *pH* is the current buffer pH value and pK_a is the apparent *pK* of that ionizable group (as listed in the database). For negatively charged groups the formula used is:

Equation 69

$$z = -\left[1 - \left(1 + 10^{(pH - pK_a)}\right)^{-1}\right]$$

Note that in the published version of the primary reference (ref. 97) the above equations were incorrectly numbered as equations 67 and 68.

The database also allows for non-titrable permanently charged groups, which are so marked by entering any negative value for the pK_a. For such entries the charge per group is taken directly as the 'charge' entry in the database.

Once the charge per group is calculated, that number is multiplied by the number of moles of that group in the protein, and then all the charges for all amino acids and attached conjugates are summed to get the net total charge, *Z*.

Caution: The actual net charge of a protein in solution may differ substantially from the theoretical values calculated by SEDNTERP. The charge calculations do not take into account monovalent ion binding, which can be significant for proteins. Furthermore, the pKas may be shifted considerably by the amino acid context (e.g. ion pairing, solvent accessibility, hydrogen bonding, *etc.*). As just one example, it is known that monoclonal antibodies usually have a much lower net charge than the calculated value (see for example ref. 75). This difference from the theoretical charge is most likely due to ion binding by the protein.

Also note that the Sample Composition form includes check boxes that determine whether the carboxy terminus and amino terminus of each polypeptide chain are included in these charge calculations.

The default amino acid and terminal group pK_a values used by the program come from Thurlkill *et al.* (2006), Ref. 93. (Note that for SEDNTERP version 1 and 2 the pK_a values were taken from Ref.

66). The amino acid pK_a values can be changed via the Estimating Databases menu.
The pK_a values for the terminal groups can be temporarily changed by altering the values shown in the text boxes seen in the image above, or permanently changed using the **Change Program Defaults...** command on the File Menu.

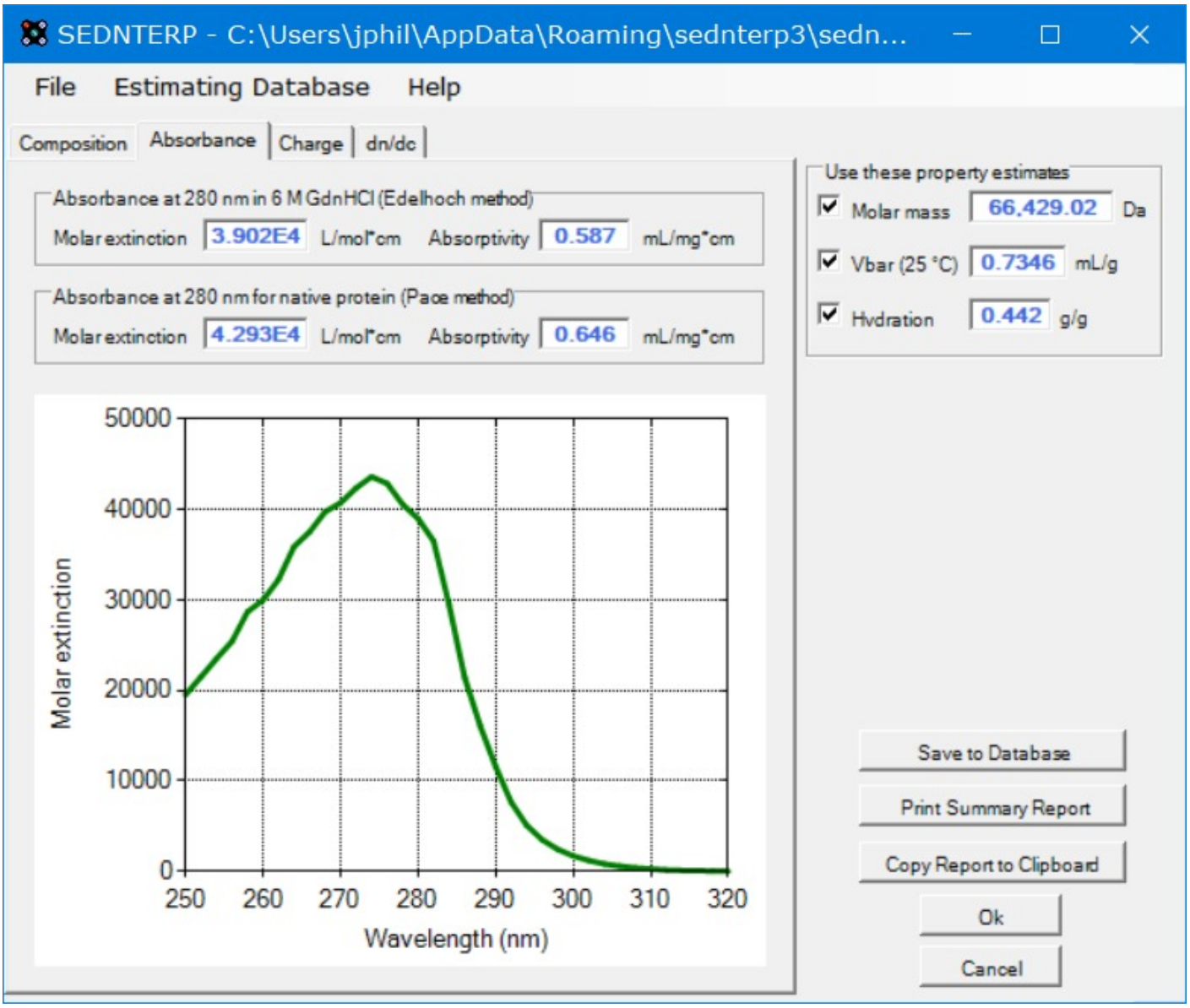
Note: Although SEDNTERP can estimate the state of ionization for each ionizable group in a protein, the sequence molar mass that SEDNTERP computes does not reflect the small mass changes associated with loss or gain of hydrogen atoms by those ionizable groups as a function of pH. That is, the sequence mass (as well the partial specific volume and *dn/dc*) are assumed to be independent of pH.

Isoelectric Point Calculation

The isoelectric point (the pH where the estimated total charge is zero) is determined using the total charge calculation procedure above, which feeds into a Van Wijngaarden-Dekker-Brent root finder to find the pH giving zero charge (to a precision of 0.00001 pH unit).

3.3.22 Calculating UV Absorbance Properties

SEDNTERP will calculate the estimated 280 nm molar extinction coefficient and absorptivity (absorbance per mg/mL per cm) of a protein, based on the composition entered on the Sample Composition form. It can also display a theoretical extinction spectrum (graph of molar extinction *vs.* wavelength), over the range from 250 to 320 nm. These calculations are shown when the "Absorbance" tab on that form is selected, as shown in the image below.



The 280 nm molar extinction and absorptivity are calculated and displayed from two different methods, the "Edelhoch method" and the "Pace *et al.* method".

The Edelhoch method uses molar extinction values for tryptophan, tyrosine, and disulfide that were measured using peptides dissolved in 6 M guanidine hydrochloride. Those measurements were made by H Edelhoch in 1967 (Ref. 63), while the method itself was described by Gill & von Hippel in 1989 (Ref. 64). That is, this method is really calculating the extinction coefficient for a fully-denatured protein in 6 M guanidine hydrochloride. However the difference in 280 nm absorbance between a native, fully-folded protein and that protein when fully-denatured is usually small (< 2%), except perhaps for proteins whose absorbance is dominated by tyrosine.

The Pace *et al.* (1995) method is, in contrast, intended to provide the extinction coefficient for a

folded, native protein (Ref. 65). It is based on using experimental extinction coefficients from the literature for 18 native proteins, and comparing their absorbance in the native state to that after unfolding in either 6 M guanidine hydrochloride or 8 M urea, as well as on 116 extinction coefficient values from the literature (for 80 different proteins), in order to calculate "best" native molar extinction coefficients for tryptophan, tyrosine, and disulfide. The "native" values so derived differ slightly from those for the Edelhoch method for tryptophan and disulfide, but the difference is largest for tyrosine.

Extinction spectra

Extinction spectra for proteins can be calculated and displayed over the wavelength range from 250 to 320 nm. Those calculations include contributions from phenylalanine residues as well as tryptophan, tyrosine, and disulfides. The spectral data used for these calculations come from Reference 67.

These spectra should also apply to conjugated proteins if the conjugates do not absorb significantly at those wavelengths. If you are working with conjugates that do absorb significantly in this region, it is possible for the user to add extinction *vs.* wavelength data to the database for existing or new conjugates, and then the conjugate contributions will also be included in the graph. (If any conjugate has extinction data over a narrower wavelength range than 250 to 320 nm, then the spectra would only be calculated and shown for that narrower range).

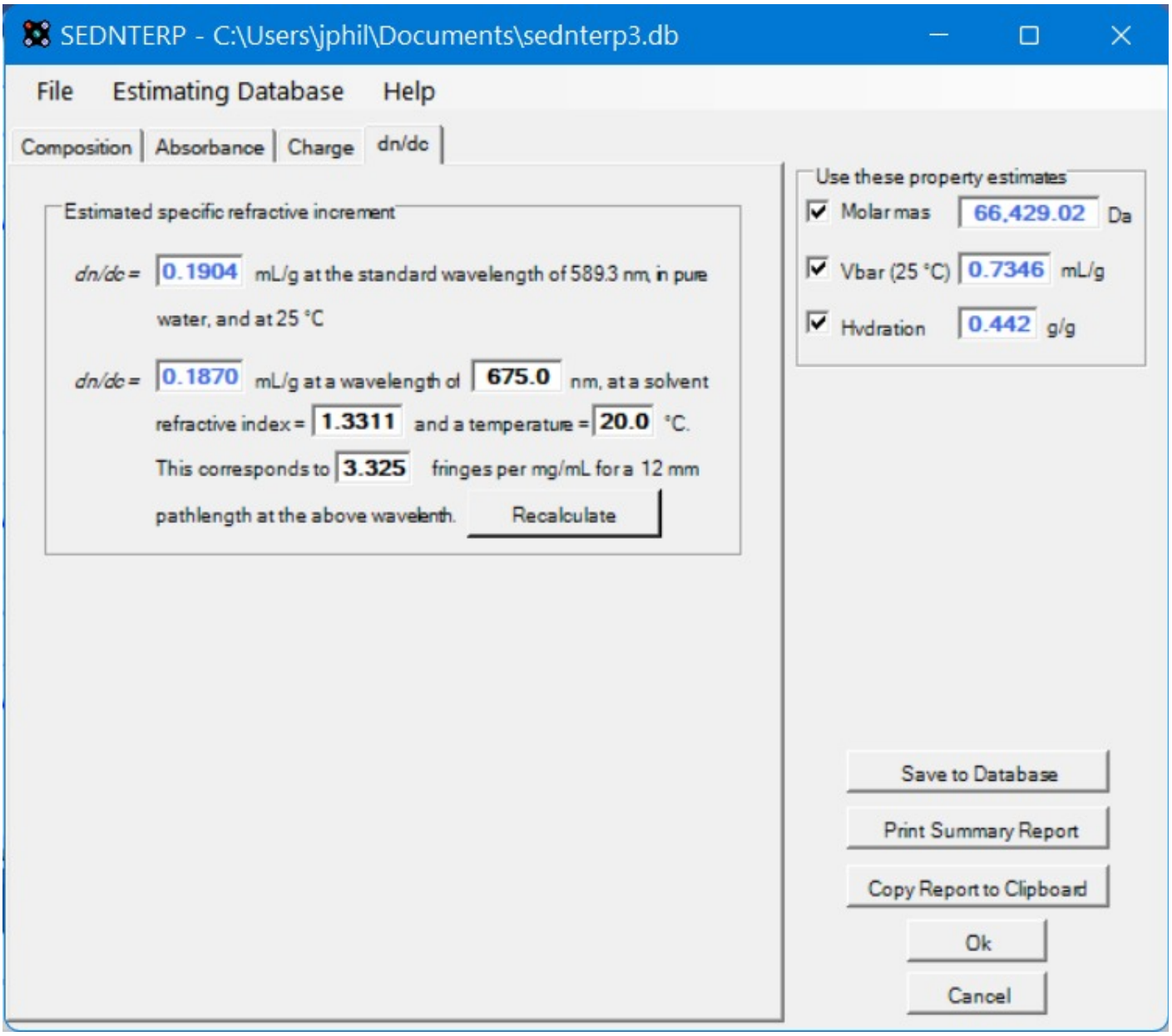
TIP: While these estimated extinction spectra may not be highly accurate for native proteins, experience shows that they often do predict whether the maximum absorbance for the native protein occurs at 280 nm, and thus predict what wavelength should be used in AUC experiments to be at the top of the absorbance peak (and thus avoid potential artifacts due to wavelength shifts).

3.3.23 Calculating Refractive Increment Estimates

A macromolecule's specific refractive increment, *dn/dc*, plays a major role in the calibration of light scattering studies (Refs. 98, 99). It also indirectly influences AUC experiments using the Rayleigh interferometer, since differences in *dn/dc* change the conversion from signal units (fringes) to weight concentration units. In other words, the *dn/dc* value determines the concentration calibration for Rayleigh interference data.

SEDNTERP can now estimate *dn/dc* values for proteins from their composition, and can also adjust those values to account for differences in buffer refractive index and/or measurement wavelength.

These values are displayed on the 'dn/dc' tab of the Sample Composition form, as illustrated below, but are not used in any other SEDNTERP calculations. The calculation method follows that described by Zhao *et al.* 2011 (Ref. 70).



Do dn/dc values really vary significantly from one protein to another, or with buffer conditions?

For the first 15-20 years after multiangle static light scattering detectors began being used for proteins on-line with size-exclusion chromatography (SEC-MALS) in the mid-1990's, it was common practice to assume that all proteins have the same value for *dn/dc* (Ref. 98). That was a reasonable

approximation for purposes of determining the stoichiometry of protein oligomers or protein complexes, and given the signal/noise ratio of the detectors in that era. However over time the detectors have become more sensitive, and the applications of SEC-MALS have become more sophisticated, including methods for conjugated proteins that distinguish the molar mass of the polypeptide portion from that of conjugates such the carbohydrate in glycoproteins or the PEG conjugated to PEGylated proteins (Ref. 99).

Such mathematical separation of components relies on differences in their *dn/dc* values and UV absorption properties, making accurate *dn/dc* values much more important. Then in 2011 Zhao *et al.* (Ref. 70) applied the approach of McMeekin et al (1964) (ref. 88) to calculate the *dn/dc* of all known and predicted proteins in the human genome, based on their amino acid composition and measured molar refractivities for the 20 types of amino acid residues. Their calculations gave a distribution of *dn/dc* values with a relative standard deviation of 1.6%, but with a predicted total range from 8.9% smaller to 13.2% larger than the mean value. Thus, their calculations show that the difference from one protein to another is indeed generally small, but sometimes can be quite significant.

It also has been common practice in both the AUC and light scattering communities to assume that *dn/dc* is the same for typical aqueous buffers, *i.e.* independent of any differences in buffer refractive index. However, this may not be a good assumption for the elution buffers used with SEC-MALS, where it is often necessary to add salts or amino acids at concentrations of 0.1-0.5 M, or alcohols at ~5% by volume, to reduce binding of the protein to the column matrix (Ref. 100). Using bovine serum albumin as an example, adding 0.2 M sodium chloride to a typical physiological buffer like Dulbecco's phosphate-buffered saline should decrease its *dn/dc* by 1.4%, a difference which is comparable to the reported 1.6% standard deviation across the human genome (Ref. 70).

Calculating the molar refractivity

The first step in calculating *dn/dc* for a protein is to calculate its refractivity per gram, R_w, as the weighted sum of values for each different amino acid residue or conjugate species, in a manner exactly analogous to calculating $\bar{v}$ from the composition. For each amino acid residue or conjugate the program database stores a molar refractivity *R* (at 25 ^{o}C and a wavelength of 589.26 nm).

Assuming refractivity is additive on a weight basis, we then can calculate R_w using

Equation 63:

$$R_w = \frac{\sum n_i R_i + n_{chains} R_{term}}{\sum n_i M_i + n_{chains} M_{water}}$$

where n_i is the number of moles of each amino acid residue or conjugate, M_i is its molar mass, R_i is its molar refractivity, n_{chains} is the number of separate amino acid chains represented in the current amino acid composition, R_{term} is the total molar refraction of a pair of terminal groups, and M_{water} is the total molar mass of a pair of terminal groups.

Note: The terms in Eq. 51 for the refraction and mass of the polypeptide N- and C-terminii (the right-hand terms in numerator and denominator) are only included in this calculation when the **Include terminal groups checkbox** on the Composition tab of the Sample Composition form is checked. When the terminal groups are added to the *dn/dc* calculation, for consistency they are also included in the $\bar{v}$ calculation (because that calculation depends on $\bar{v}$).

The residue refractivity values in the program database are from McMeekin *et al.* 1964 (ref. 88). Note that in Zhao *et al.* 2011 (ref. 70) the value listed for cysteine is actually that for cystine. For the conjugates, the refractivity values were taken from the literature where available, or calculated from *dn/dc* values in the literature, or when no experimental data were found the values predicted by the ACD/Labs Percepta platform (as available through ChemSpider.com) were used when available.

Calculating the protein's refractive index

Next, using the protein's partial specific volume $\bar{v}$ and the Lorentz-Lorenz formula, $R_w = \bar{v}(n^2 - 1)(n^2 + 2)$, we can then solve for the protein's refractive index, n_p, using

Equation 64:

$$n_p = \sqrt{\frac{2R_w + \bar{v}}{\bar{v} - R_w}}$$

Calculating *dn/dc*

Finally, using a formula for dilute solutions from Weiner (ref. 86) we can calculate *dn/dc* (at 25 ^{o}C and 589.26 nm) via

Equation 65:

$$dn/dc = \frac{3}{2}\bar{v}n_b\frac{n_p^2 - n_b^2}{n_p^2 + 2n_b^2}$$

where n_b is the refractive index of the buffer.

Correcting *dn/dc* for wavelength and temperature

When needed, the estimated *dn/dc* values are adjusted for wavelength using the formula of Perlmann & Longsworth 1948 (ref. 87)

Equation 66:

$$(dn/dc)_\lambda = (dn/dc)_{578} \times \left(0.940 + \frac{20{,}000\ \text{nm}^2}{\lambda^2}\right)$$

When needed, the estimated *dn/dc* values are adjusted for temperature using a formula from Zhao *et al.* 2011 (ref. 70)

Equation 67:

$$(dn/dc)_T = (dn/dc)_{25} \times \left[1 + (25 - T)\frac{0.0025}{30}\right]$$

How accurate are these dn/dc estimates?

The accuracy of these *dn/dc* estimates is difficult to assess with any certainty. As pointed out in ref. 70, the largest source of error is likely to be introduced from errors in $\bar{v}$, and it is known that errors in $\bar{v}$ are larger for small peptides than for larger proteins. Furthermore, it also seems unclear whether $\bar{v}$ necessarily accurately reflects the molecular volume that is relevant for calculating the protein's refractive index and *dn/dc*.

On the other hand, this author has compared these estimated *dn/dc* values with 4-5 values that were determined in his lab via static light scattering (back-calculating *dn/dc* from the known protein molar masses). In every case the estimated *dn/dc* was within 1% of the experimental value. That is excellent agreement considering that (1) the absolute calibration of the light scattering and refractive index detectors is likely uncertain by ~1%, and (2) some of these experiments were in buffers whose refractive index (as calculated by SEDNTERP from the buffer composition) differs significantly from pure water.

4 Program elements, forms, and features

4.1 Program Strategy & Layout

Program Strategy

The overall strategy or goal of the SEDNTERP program is to allow the user to enter information and complete the desired calculations quickly and easily.

The program will default to approximate values for all parameters that admit even any guess. Thus, by entering just a few parameters directly from the experiment, a reasonable estimate of the desired molecular parameters can be calculated by the program.

For example, the program will assume that the buffer density is that of pure water. Not many relevant biological molecules will even dissolve in water, and even if they did pure water is definitely not a suitable solvent for AUC experiments (you must have enough ionic strength to suppress solution non-ideality), so this guess is almost always wrong! Nonetheless since many dilute buffers have densities close to that for water, this guess is reasonable for quick calculations.

On the other hand, SEDNTERP also allows the user to calculate or enter almost every value used in the calculations. Entering only 2-3 numbers (known molar mass + s*, or sigma + RPM + s*) will allow the program to calculate reasonable results for the molecule's hydrodynamic properties. On the other hand, it is possible to enter over 24 parameters that affect the calculations, including the individual buffer components. It is even possible to add new buffer components to the database so that SEDNTERP can calculate the density, viscosity, and refractive index of any buffer for which the experimenter has the proper information. Of course, the entering of more information will increase the accuracy of the calculations performed and of your interpretation of the results.

Furthermore, the SEDNTERP program has the strategy to follow flexible paths of computation. The exact order of calculations depends on the sequence of the user's input of information.

Color codes indicate where a particular value comes from

Because some parameters can be either calculated values or user-input values, SEDNTERP color-codes the calculated values as **blue text**, the user-input values as **black text**, and user-input 'overrides' of values which are normally calculated as **green text**.

For example, if the amino acid sequence of a protein sample is entered, SEDNTERP will calculate and use the molar mass by sequence for further calculations. However the user can choose to ignore this calculation and instead enter sedimentation equilibrium data, or directly enter a known molar mass value (*e.g.* from mass spectroscopy). The molar mass by sequence is generally more

accurate than other computed masses, yet if the protein undergoes post-translational modification, its mass may have to be determined by experiment and not by sequence. SEDNTERP would allow you to estimate the vbar of the protein by sequence and still use a molar mass determined by sedimentation equilibrium or mass spectroscopy.

The user-supplied information (and many calculated results) can be saved into the program database

SEDNTERP allows the user to save all the information entered by the user in separable tables within a unified database. In this way, the details of the analysis can be saved and reproduced later (*e.g.* to document calculations used in a publication). However the user-supplied information can also be modified easily at a later date, and by saving into the database the user can avoid needing to re-enter information that applies to multiple experiments (such as common samples or buffer compositions).

Default values

SEDNTERP uses default values for many parameters that may be entered by the user, which allows many quantities to be at least roughly estimated without requiring that the user knows and enters every detail. Note however that reliance on these default values will in most cases limit the accuracy of the calculations performed by SEDNTERP.

These default values can be changed using the 'Change Program Defaults' item on the File menu of the main form.

Details about each default value

SEDNTERP automatically enters default values for the following parameters (standard default values are in parenthesis; default values currently in use may be different than those listed here due to changes by previous use):

1. The temperature of the sedimentation equilibrium experiment (25 ^{O}C)
2. The temperature of the sedimentation velocity experiment (20 ^{O}C)
3. The temperature of the diffusion experiment (25 ^{O}C)
4. The $\bar{\nu}$ of the sample (0.73 mL/g, a good guess for globular proteins)
5. The density of the buffer (0.998213, the density of pure water at 20 ^{O}C)
6. The viscosity of the buffer (0.010016 poise, the viscosity of pure water at 20 ^{O}C)
7. The refractive index of the buffer (1.333359 at 20 ^{O}C and 589.26 nm, the value for pure water)
8. The pH of the buffer (7.0)
9. The hydration, δ_1 (0.3 g/g, a commonly-used 'typical' value for globular proteins)

10. The default pK_a values for the amino terminus and carboxyl terminus (8.00 and 3.67, respectively)
11. The default density for the bound water of hydration (0.998213, the density of pure water at 20 ^{O}C)
12. The concentration dependence of the sedimentation coefficient, k_s (9 mL/g, a good guess for compact globular proteins)
13. The concentration dependence of the diffusion coefficient, k_D (0.9 mL/g, a 'not unreasonable' guess for compact globular proteins at physiological ionic strength, when measured by methods involving a concentration gradient). This default value is definitely the most uncertain, and the rationale for choosing this value can be found here.
14. The default wavelength to use when calculating the refractive index (589.26 nm)
15. The default wavelength to use for calculating the *dn/dc* value used in experiments (675 nm)
16. The default method for computing the Equivalent Radius, either the Teller method or the vbar method (Teller)
17. The default relative size for the program forms, as a percentage of the nominal size (100)
18. The default value for the covolume term when calculating the partial specific volume (12.4 mL/mole)
19. The default total molar volume for the amino- and carboxy-terminal groups (-6.9 mL)
20. The default total molar refractivity of the amino- and caroboxy-terminal groups (3.73)

If you commonly use a particular buffer for your experiments, you may wish to reset the default density, viscosity, refractive index, and pH to correspond to that buffer.

A Main Form Separated into Category Tabs Leads to Other Forms or Dialog Boxes

One strategy of SEDNTERP is to try to separate information about the sample, buffer, experiment(s), and computed hydrodynamic results onto different 'tabs' of the main form. Such a separation, however, cannot be complete, so some property choices and inputs may appear on more than one tab. Furthermore, some sample properties may vary significantly depending on the buffer conditions or even the experiment temperature, so information entered on different tabs may be interrelated.

Additional optional information is entered onto separate forms, such as the Sample Composition form, the Buffer Composition form, the Evaluate ks and Extrapolate s0(20,w) form, and the Saving Data in Databases form.

4.2 Description of the Databases

SEDNTERP uses a single SQLite database file (usually the file sednterp.db). SQLite is public-domain

database that can run on virtually any computer platform. The SEDNTERP database is a "relational" database where certain tables may contain links pointing to data located in other tables. Although the database is contained in a single file, it can be conceptually divided into two parts, each of which contains multiple tables, as detailed below.

The "User" or "Experimental" database

The "User" or "Experimental" portion of the database contains tables that store information about:

1. samples
2. buffers
3. experiments (three different tables for sedimentation velocity, sedimentation equilibrium, or diffusion experiments)
4. computed hydrodynamic results
5. dilution series (measurements of the sedimentation or diffusion coefficient of the same sample in the same buffer at multiple concentrations)

Each table contains multiple rows ("records" in database terminology), with one row for each different sample, buffer, etc. By storing this information in the database, the user can avoid repetitively re-entering information that applies to many cases. For example, if a set of experiments uses the same buffer, the same saved Buffer record may be used for all of them. Similarly, if the same sample is tested in many buffers, the same Sample record may be used. If an experiment needs to re-examined, or a hydrodynamic result needs to be documented and reproducible because it is going to be published, those records can be recalled. Further, if different assumptions are used in analyzing an experiment, the different results can each be saved as a separate Results record.

When the program is first installed this database contains just a few records intended as examples (some of which contain real published results obtained by John Philo in the 1990's). The user can then add to this default database, and can later remove any of the start-up records if desired (from the Saving Data in Databases form). The records in the Experimental database are updated either through items on the File Menu of the main form, or through dedicated buttons on the Sample Properties major tab, the Buffer Properties major tab, the Sedimentation Velocity minor tab, the Sedimentation Equilibrium minor tab, the Diffusion minor tab, the Derived Hydrodynamic Results major tab, the Sample Composition form, the Buffer Composition form, the Evaluate ks and Extrapolate s0(20,w) form, or the Evaluate kD and Extrapolate D0(20,w) form.

This 'User' portion of the database contains 7 different major types of records:

- 'Sample' records contain information about the sample's molar mass, $\bar{\nu}$, hydration, k_s, and k_D, as well as variables indicating where those sample properties come from (*e.g.* calculated

from composition or directly entered as a known value). They also record whether the sample was denatured using a defined denaturant, the pK_a values for the amino and carboxy terminii, the samples oligomeric state, and optionally the sample's composition (amino acids + conjugates). Sample records may also contain links to sedimentation or diffusion dilution series records.

- 'Buffer' records record the buffer density, viscosity, and refractive index (and again variables indicating where those values come from), the pH, the fractions for heavy isotopes of water (if any), and optionally the buffer composition.
- 'Velocity Experiment' records have information about the experiment temperature and date, the sample's sedimentation coefficient (both s^* and $s_{20,w}$ or $s^0_{20,w}$), the sample's diffusion coefficient (if determined, both D^* and $D_{20,w}$ or $D^0_{20,w}$), its molar mass, the sample concentration (optional), the buoyant molar mass, Stokes radius, k_s and k_D, the sample $\bar{\nu}$ and buffer density and viscosity at the experiment temperature. In addition, each Velocity Experiment record usually is linked to both a Sample record and a Buffer record, and may optionally be linked to an Equilibrium Experiment, Diffusion Experiment, Sedimentation Dilution Series, and/or Diffusion Dilution Series record. Linking the velocity experiment to an equilibrium experiment would normally be done when the molar mass used to interpret the velocity experiment results comes from that equilibrium experiment. Linking the velocity experiment to a diffusion experiment would typically be done when the sedimentation coefficient is being combined with a diffusion experiment to determine the molar mass.
- 'Equilibrium Experiment' records contain information about the experiment date, the sample's experimental molar mass as determined by sedimentation equilibrium. That information can be entered as either a true molar mass, a buoyant molar mass, or a reduced apparent molar mass (σ) value. It also contains the experiment temperature, the sample $\bar{\nu}$ and buffer density and viscosity at the experiment temperature, and the rotor speed (required only for σ mass units). In addition, each Equilibrium Experiment record usually is linked to both a Sample record and a Buffer record, and may optionally be linked to a Velocity Experiment or Diffusion Experiment record (which is typically done when the molar mass used for interpreting those other experiments comes from the equilibrium experiment).
- 'Diffusion Experiment' records contain the experiment date, either the measured diffusion coefficient (both D^* and $D_{20,w}$ or $D^0_{20,w}$) or the measured Stokes radius (depending on which was returned by the instrument), the experiment temperature, the buffer viscosity at that temperature, the sample concentration (optional), the k_D value and where it came from, and the sample molar mass (if that was determined by combining sedimentation and diffusion experiments). In addition, each Diffusion Experiment record usually is linked to both a Sample record and a Buffer record, and may optionally be linked to a Velocity Experiment record (which is typically done when the molar mass has been calculated by combining the sedimentation and diffusion experiments), or to a Sedimentation Equilibrium

Experiment record (which is typically done when the molar mass used for interpreting the other experiment comes from the equilibrium experiment), or to a Diffusion Dilution Series record.

- 'Hydrodynamic Result' records record hydrodynamic properties that have been calculated from a particular Sedimentation Velocity or Diffusion Experiment record, such as the experimental standardized frictional coefficient, the equivalent radius and minimum frictional coefficient, as well as information about hydration and asymmetry (and which shape model was used for the asymmetry calculations). Each Result record usually is linked to Sample record, either a Velocity or Diffusion Experiment record, and possibly an Equilibrium Experiment record (if the molar mass comes from sedimentation equilibrium).
- 'Sedimentation Dilution Series' or 'Diffusion Dilution Series' records contain multiple measurements of the sedimentation or diffusion coefficient of a particular sample at a series of different concentration, and optionally the uncertainty of each experimental measurement, for purposes of extrapolating the sedimentation or diffusion coefficient to zero concentration (to determine $s^0_{20,w}$ or $D^0_{20,w}$), and/or to determine k_s or k_D. These records also store the r.m.s. deviation of the data points from the best fitted line, and whether that fit assumed that all data points had equal uncertainty (equal weight).

For the User database portion, it is obviously important for each user to create an intelligent naming scheme if many experiments, samples, and/or buffers are to be stored in the database. Also, it is may be useful for different users from the same lab to have their experiments in a totally different database file in order to avoid overwriting data records because of repetitive names or human error.

Note that it is possible for multiple users to share a single database if that database is stored in a location that all users can access (for example on a shared network drive). However if a network drive is used, it is important that the path to the database file remains fixed, which may require that the network drive is a "mapped" drive (it is assigned to a specific drive letter). The location of the database file that is used when the program starts can be changed by opening that file using the **File...Open New Database** menu item.

The Physical Constants database

The 'Physical Constants' portion of the database contains essential information concerning:

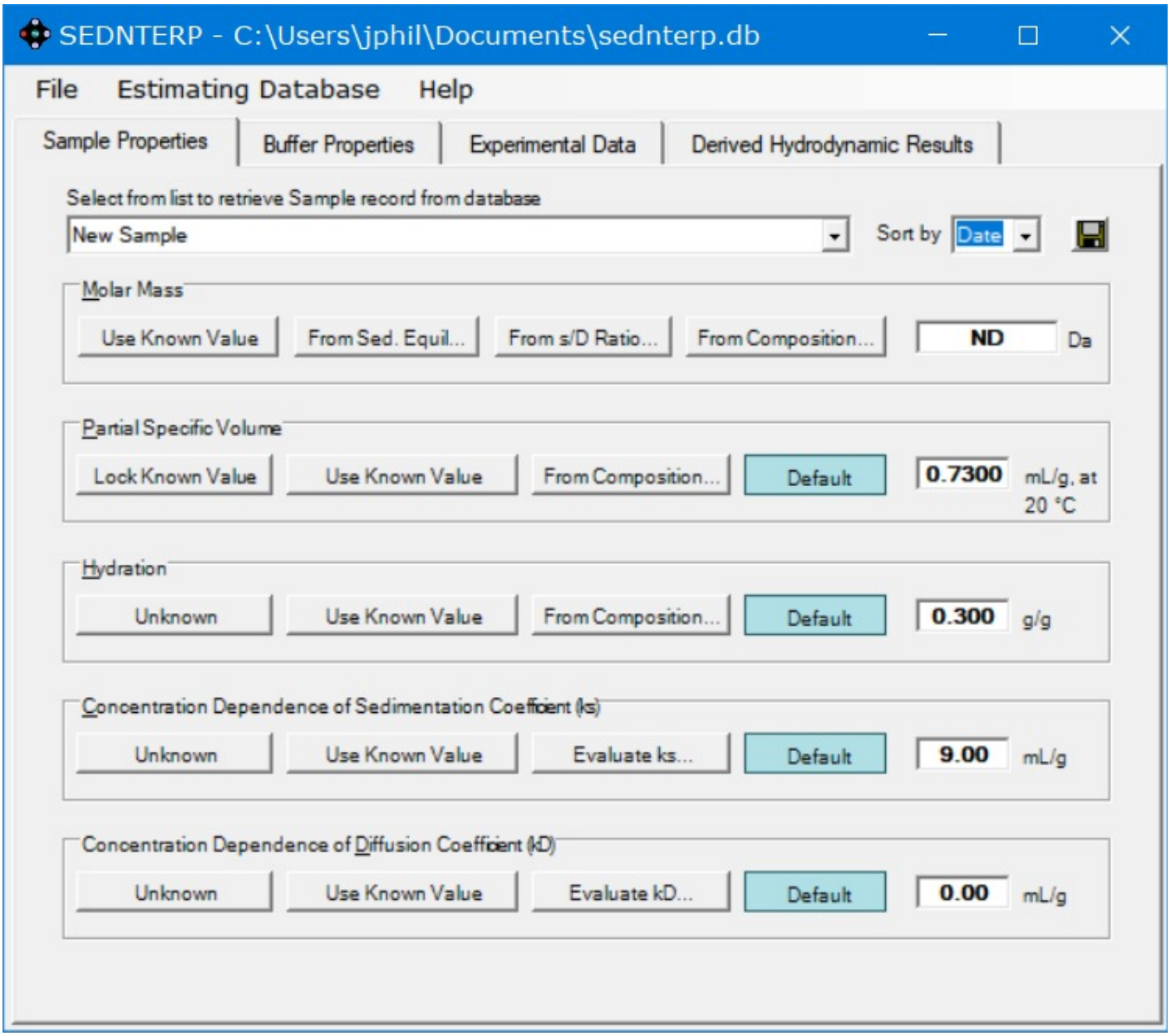
- Properties of the amino acids and conjugates:
  - molar mass

    - $\bar{v}$
    - hydration (at pH <= 4, and at pH > 4)
    - electric charge and pK_a (if any)
    - molar extinction at 280 nm
    - molar refractivity
- Data about potential buffer components (ingredients):
    - molar mass
    - density (if a liquid)
    - coefficients to calculate how this component changes the buffer density as a function of concentration
    - coefficients to calculate how this component changes the buffer viscosity as a function of concentration
    - coefficients to calculate how this component changes the buffer refractive index as a function of concentration
- Information about 'defined denaturants' used to calculate $\bar{v}$ for denatured proteins

All of these tables can be added to or updated through the Estimating Database menu on the main form. More information about editing these tables can be found in the 'Physical Properties Tables' section of this document.

4.3 Main Data Entry form

The image below shows a new, empty Main Data Entry form (as it looks when the program starts).



This main form is divided into 4 pages or sections using a 'tab' control with its tabs located at the top, just below the menu bar. The 'Sample Properties' section is shown above, and the 'Buffer Properties', 'Experimental Data', and 'Derived Hydrodynamic Results' sections are each accessed by clicking on the corresponding tab near the top. The Experimental Data 'major' tab page is itself divided into 3 'minor' tab pages for different types of experiments (sedimentation velocity, sedimentation equilibrium, or diffusion). This separation of information into different categories is intended to make things easier and clearer for the user. However, this separation cannot be complete, so some property choices and inputs may appear on more than one tab. Furthermore, the values for some sample properties and/or experimental values such as sedimentation coefficients may vary depending on information found on other tabs (such as buffer

pH and composition).

It should also be noted that the values for some properties or parameters may not be important for all situations and applications. For example, the entries for the concentration dependence of the sedimentation and diffusion coefficients shown above concern solution non-ideality effects that are not important if the experiments are run at sufficiently low concentrations.

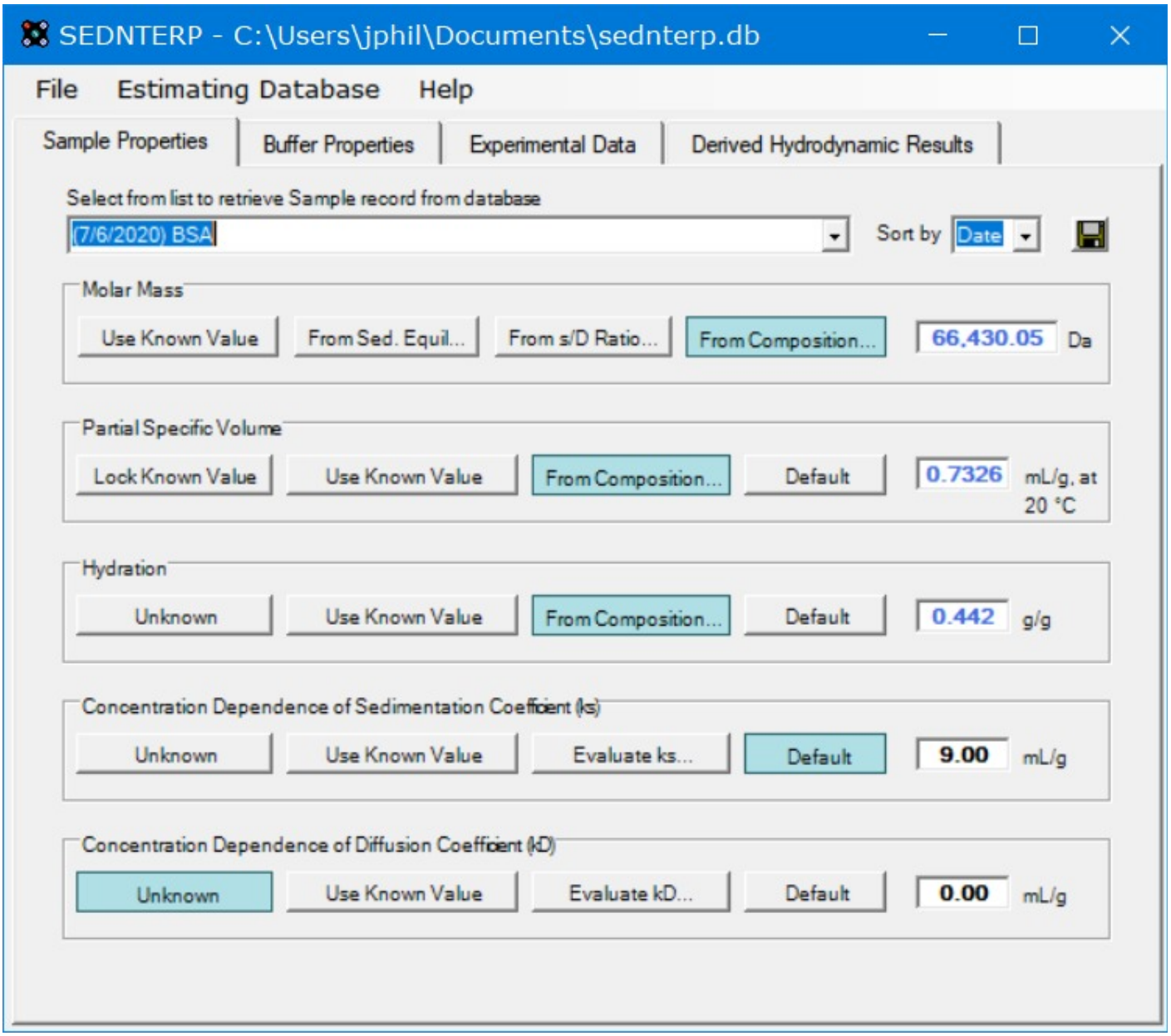
An important feature of SEDNTERP is that many of the properties may come from different sources (for example an experimental value *vs.* a computed or interpolated value), and nearly all user inputs also have a 'default' value (which may be 'close enough', at least for initial or rough calculations). Therefore it is often important for SEDNTERP to visually indicate which type or source of information is currently in use. In the image above, the light blue background of the four **Default buttons** signals that each of those properties is currently set at its default value.

Another important feature of SEDNTERP is that the information the user enters can be saved to, and later restored from, the program database. Each of these four major tab pages of the main form also corresponds (at least approximately) to a different type of table in the database (one table for sample properties, one for buffer properties, one for derived hydrodynamic results, and one table each for the 3 different types of experiments). Each of these major or minor tabs has a pull-down list near the top (like the one shown above near the top of the Sample Properties tab) that contains a list of previously-stored data records that can be re-loaded from the database. Similarly, each of these major and minor tabs has a small **Save button** (with a disk icon) that initiates saving new records to the database, similar to the one seen at top right on the image above.

Further details about the properties and controls on each of the four major tabs of the can be found in their individual topic pages: Sample Properties major tab, Buffer Properties major tab, Experimental Data major tab, and Derived Hydrodynamic Results major tab.

4.4 Sample Properties major tab

The Sample Properties major tab of the Main form displays information about properties of a particular sample (usually a protein or other macromolecule) that is the subject of an AUC and/or diffusion experiment. The image below shows what this tab page looks like after the BSA sample record (included in the start-up database) has been loaded.



This tab page contains 5 sections, each about a different sample property:

- molar mass
- partial specific volume ($\bar{\nu}$)
- hydration (δ_1)
- concentration dependence of the sedimentation coefficient, k_s
- concentration dependence of the sedimentation coefficient, k_D

For each of those five properties, the source of the property value used by SEDNTERP in its

calculations is indicated by which of the 4 buttons within that panel shows the light blue background.

Molar mass panel

The **Use Known Value** button enables the user to enter any value for the molar mass directly into the text box on the right. For example, this could be a value determined by mass spectroscopy. When the button is clicked the cursor is placed into the text box for data entry.

The **From Sed. Equil...** button indicates the molar mass value is coming from a sedimentation equilibrium experiment. When this button is clicked the user will be taken to the Sedimentation Equilibrium minor tab of the Main form so the information about that equilibrium experiment can be entered.

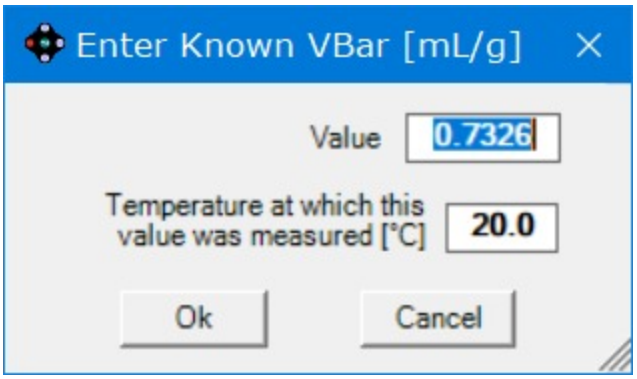
The **From *s/D* Ratio...** button indicates that the molar mass has been determined from the ratio of the sedimentation and diffusion coefficients. That ratio could come directly from a single sedimentation velocity experiment where the data analysis returned both *s* and *D* (or *s* and *M*), or it could come from separate sedimentation and diffusion experiments. When this button is clicked, if no sedimentation coefficient is available then the user will be taken to the Sedimentation Velocity minor tab of the Main form so the information about the sedimentation velocity experiment can be entered, whereas if no diffusion coefficient is available the user will be taken to the Diffusion minor tab.

The **From Composition...** button indicates that the molar mass has been computed from the sample's total composition (amino acids + any conjugates). When the button is clicked the 4 main tabs are temporarily replaced by the Sample Composition form so the user can enter the composition information.

Partial specific volume panel

The **Lock Known Value** button indicates that user will enter a known $\bar{\nu}$ value into the text box on the right, and also that this $\bar{\nu}$ value should not be temperature corrected by SEDNTERP (this is indicated by the 'No T correction' message that appears in red below the text box).

The **Use Known Value** button allows the user to enter a value for $\bar{\nu}$ at a particular temperature (preferably from an actual measurement you or someone else has made at some temperature, but perhaps from a standard table of physical constants). The entered value will then be corrected for any temperature difference between the entered value and the experiment. When this button is pressed it brings up a small dialog box for entering the known $\bar{\nu}$ and its measurement temperature, as shown below:



The **From Composition...** button indicates that the $\bar{\nu}$ has been computed from the sample's total composition (amino acids + any conjugates). When the button is clicked the 4 main tabs are temporarily replaced by the Sample Composition form so the user can enter the composition information.

The **Default** button sets $\bar{\nu}$ to the current default value (which is 0.73 mL/g when the program is installed, but which can be changed to whatever the user prefers.)

Hydration panel

The **Unknown** button indicates that the user knows nothing about the hydration of the sample, and wants any subsequent hydrodynamic calculations to make no assumptions about hydration (and consequently to merely calculate the extreme values consistent with the experimental data (the maximum possible hydration and the maximum possible asymmetry). When this button is pressed the hydration value displayed becomes zero.

The **Use Known Value** button enables the user to enter any positive value for δ_1 directly into the text box on the right. When the button is clicked the cursor is placed into the text box for data entry.

The **From Composition...** button indicates that the δ_1 has been computed from the sample's composition. When the button is clicked the 4 main tabs are temporarily replaced by the Sample Composition form so the user can enter the composition information.

The **Default** button sets δ_1 to the current default value (which is 0.3 g/g when the program is installed, but which can be changed to whatever the user prefers.)

Concentration dependence of *s or D* panels

The **Unknown** button indicates that the user knows nothing about the value of k_s or k_D for the sample, and therefore SEDNTERP should not attempt to make a concentration-dependence

correction. When this button is pressed the corresponding k_s or k_D value displayed becomes zero.

The **Use Known Value** button enables the user to enter any value for k_s or k_D directly into the text box on the right. When the button is clicked the cursor is placed into the text box for data entry.

The **Evaluate ks...** or **Evaluate kD...** buttons will bring up a separate form where the user can enter a series of *s* or *D* values measured at different sample concentrations, and then fit those values to a line to evaluate k_s or k_D. Note that when invoked from the Sample Properties tab this form will not return the extrapolated value at zero concentration as the $s^0_{20,w}$ or $D^0_{20,w}$ value---that can only be done when this form is invoked from a Sedimentation Velocity or Diffusion Experiment minor tab.

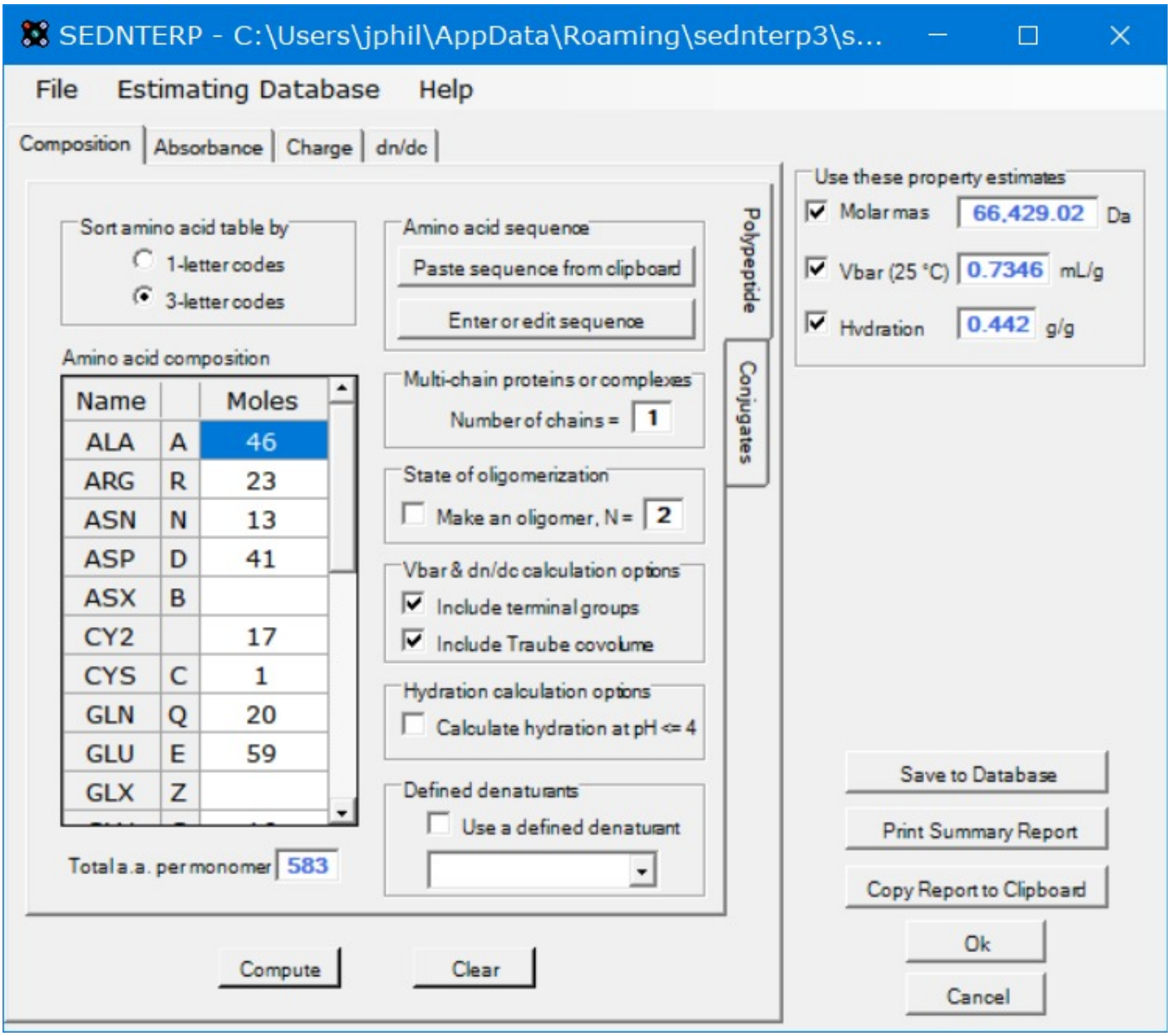
The **Default** button sets k_s or k_D to the current default value (which is 9 mL/g for k_s and 0.9 mL/g for k_D when the program is installed, but which can be changed to whatever the user prefers.)

Retrieving or saving a Sample record from/to the database

A previously saved Sample record can be retrieved from the database when chosen by name from a drop-down list box located at the upper left in this section. Whether that list is sorted by name, or by the date the Sample record was last modified, is controlled by the **Sort by** drop-down list to its right. The **Save** button at the upper right (with a disk icon) will initiate saving a new Sample record (or over-writing the current one) by bringing up the Saving Data in Databases form.

4.5 Sample Composition form

The sample composition form allows the user to enter the composition of a protein and also allows access to values for multiple sample properties that can be computed from the composition.



The image above shows the sample composition form loaded with values for BSA. At the upper left is a tab control with 4 tabs:

- The **Composition** tab, where the user enters information about the amino acid composition and any attached conjugates
- The **Absorbance** tab, where the calculated 280 nm extinction coefficient and absorptivity, as well as estimated extinction spectrum, are displayed
- The **Charge** tab, where the calculated charge, isoelectric point, and titration curve are displayed
- The **dn/dc** tab, where the calculated specific refractive increment (*dn/dc*) and

the expected fringes per mg/mL in an AUC instrument are displayed

At the upper right is a group of 3 check boxes which govern whether the calculated molar mass, $\bar{v}$, or hydration are returned to the main form and used in further calculations.

The Composition tab

The Composition tab page is itself split into two vertical tabs on its right-hand side, one about the amino acid composition of the polypeptide portion, and one about the conjugates.

As shown above, on the Polypeptide minor tab page the amino acid composition is listed in a spread-sheet like table at the upper left. The amino acids are listed in alphabetical order on the basis of either their 1-letter or 3-letter codes, as selected by a pair of radio buttons above the grid. You may switch the ordering preference without losing the entered data.

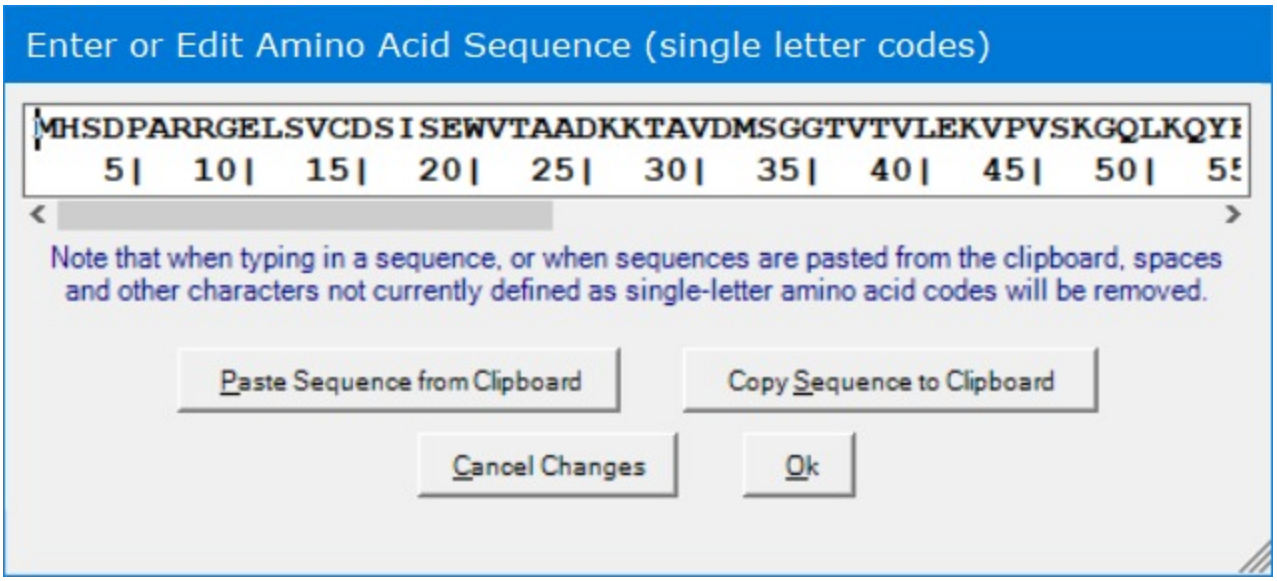
The number of moles of each amino acid can be be directly manually entered into the table. To speed data entry, computations are not made automatically during data entry to the table, but the computations are performed when the "**Compute Values**" button is pressed or when you exit the table using the Tab key. Entries are provided for the twenty common amino acids, as well as for GLX and ASX (average of the values for GLN and GLU and ASN and ASP respectively) and UNK, the average value of the twenty amino acids, for use when the residue identity is unknown. (The use of these average values for the amino acids will, of course, decrease the accuracy of subsequent calculations.)

Entering an amino acid sequence

Alternatively, the AA composition can be computed from a known amino acid sequence (which must be in 1-letter AA codes) by copying that sequence to the Windows Clipboard and then clicking on the **Paste sequence from clipboard** button.

TIP: Spaces, numbers, punctuation, *etc.* included within an AA sequence will be ignored. Thus you can directly paste in sequences copied from various databases that include spaces, sequence numbers, line breaks, *etc.*

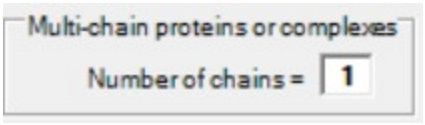
If you want to manually enter a sequence, or have an AA sequence on the Clipboard that needs to be edited before being summed (for example, to remove a leader sequence or His tag), click on the **Enter or edit sequence** button to bring up the dialog box shown below:



If you are pasting sequences copied to the clipboard, you must use the Paste command button. You should **not** use the Windows shortcut Paste key combination (CTRL-V) to directly paste into the text box holding the sequence, since this will bypass the elimination of irrelevant spaces and other potentially invalid characters. While using this editing form you can also copy the edited sequence back to the Clipboard.

After sequence entry is complete, the total numbers of each amino acid are summed and entered into the table. When this summation is complete, if the sequence contains any cysteines you are then asked how many of those cysteines should be converted into disulfides. By default SEDNTERP will suggest converting all cysteines to disulfides (all but 1 if the total is an odd number). Note that it is both possible, and sometimes appropriate, to convert an odd number of cysteines to disulfides (thus creating a half disulfide).

Handling multi-chain proteins or complexes

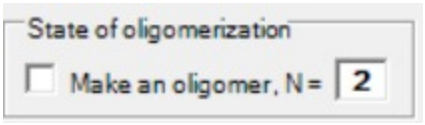


If you are working with a protein that has multiple covalently-linked polypeptide chains (like insulin), or a non-covalent multi-protein complex that you want to treat as a single entity, such situations can be handled by:

1. Concatenating together the sequences for each polypeptide chain (the order of the chains in that sequence doesn't matter), or by simply manually entering the summed amino acid composition for all the chains
2. Changing the value in the **Number of chains** textbox.

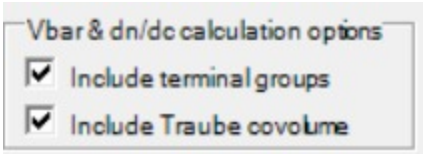
Increasing the number of chains will add additional amino- and carboxy-terminus groups into the molar mass and charge calculations, and optionally will also add them into the $\bar{\nu}$ and *dn/dc* calculations (if the **Include terminal groups** option is checked)

Other options on the Polypeptide composition tab

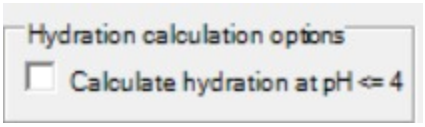


Once the amino acid composition is complete and the number of chains has been set, you may wish to make an oligomer of that entity (*e.g.* when the species being measured in your sedimentation or diffusion data is a non-dissociating oligomer). This can be done by checking the **Make an oligomer** check box and then entering the appropriate oligomer stoichiometry in the text box to the right.

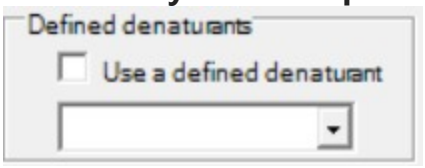
NOTE: To illustrate the usage of the oligomer and multi-chain options, for a monoclonal antibody you would combine the sequences for light chain and heavy chain (one after the other, in either order), and after pasting in that combined sequence you would set the number of chains at 2, check **Make an oligomer**, and enter 2 for the state of oligomerization, thus creating the typical immunoglobulin $(HL)_2$ hetero-oligomer.



In versions 1 and 2 of SEDNTERP the partial specific volume was calculated using the method of Cohn and Edsall (Refs. 4, 5, 6, 40). However that method neglects 2 relatively small terms, which can now be included if the corresponding check boxes are checked. Checking **Include terminal groups** will add the volume associated with the N- and C- terminii into the $\bar{\nu}$ calculation, and similarly will also add the refraction of the terminal groups into the *dn/dc* calculation. Checking **Include Traube covolume** will add a small volume correction related to packing efficiency into the $\bar{\nu}$ calculation. For fairly large proteins (>50 kDa) the effect of adding these terms is quite small (a few digits in the 4th decimal place), but for small peptides the use of these additional terms may significantly improve the accuracy. See Calculating Vbar Estimates and Calculating Refractive Increment Estimates for further details about these options.



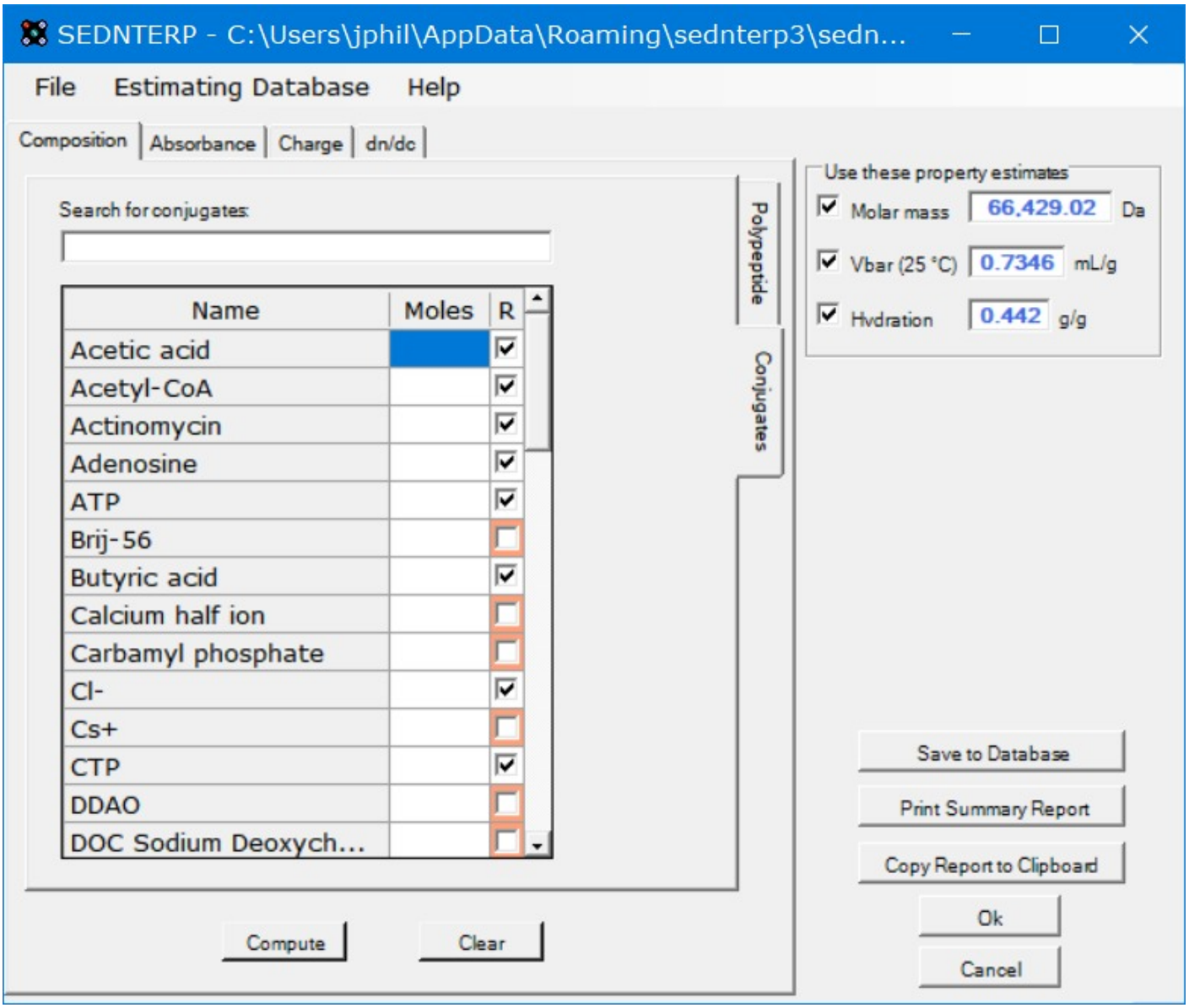
It is important to note that **the hydration estimate calculation depends on pH**, with different values being returned for pH values <= 4 (see details here). Therefore you may need to check the **Calculate hydration at pH <= 4** check box.



Special defined denaturants affect the vbar estimates provided by SEDNTERP. The equations used for such computations are discussed in the Calculating Vbar Estimates topic.

Adding conjugates

Clicking the **Conjugates** minor tab brings up a table showing a table of known conjugates (shown below), in which you can manually enter the number of moles for each potential conjugate species. (Note that the number of moles does not have to an integer).



The Search box above the conjugates table allows you to quickly search the list for a particular name.

NOTE: The conjugates table no longer includes the 'non-covalent' conjugates that were present in SEDNTERP versions 1 and 2. Those 'loosely-bound' conjugates had created significant confusion. It is certainly true that a compound may be weakly bound and its binding site(s) may be only partially occupied under given conditions. However for the purposes of this program we must consider binding to be binary: the

compound is either bound, or not bound. If it is bound, it contributes to the calculated molar mass, partial specific volume, *etc.*, and whether it is bound covalently or non-covalently is really irrelevant.

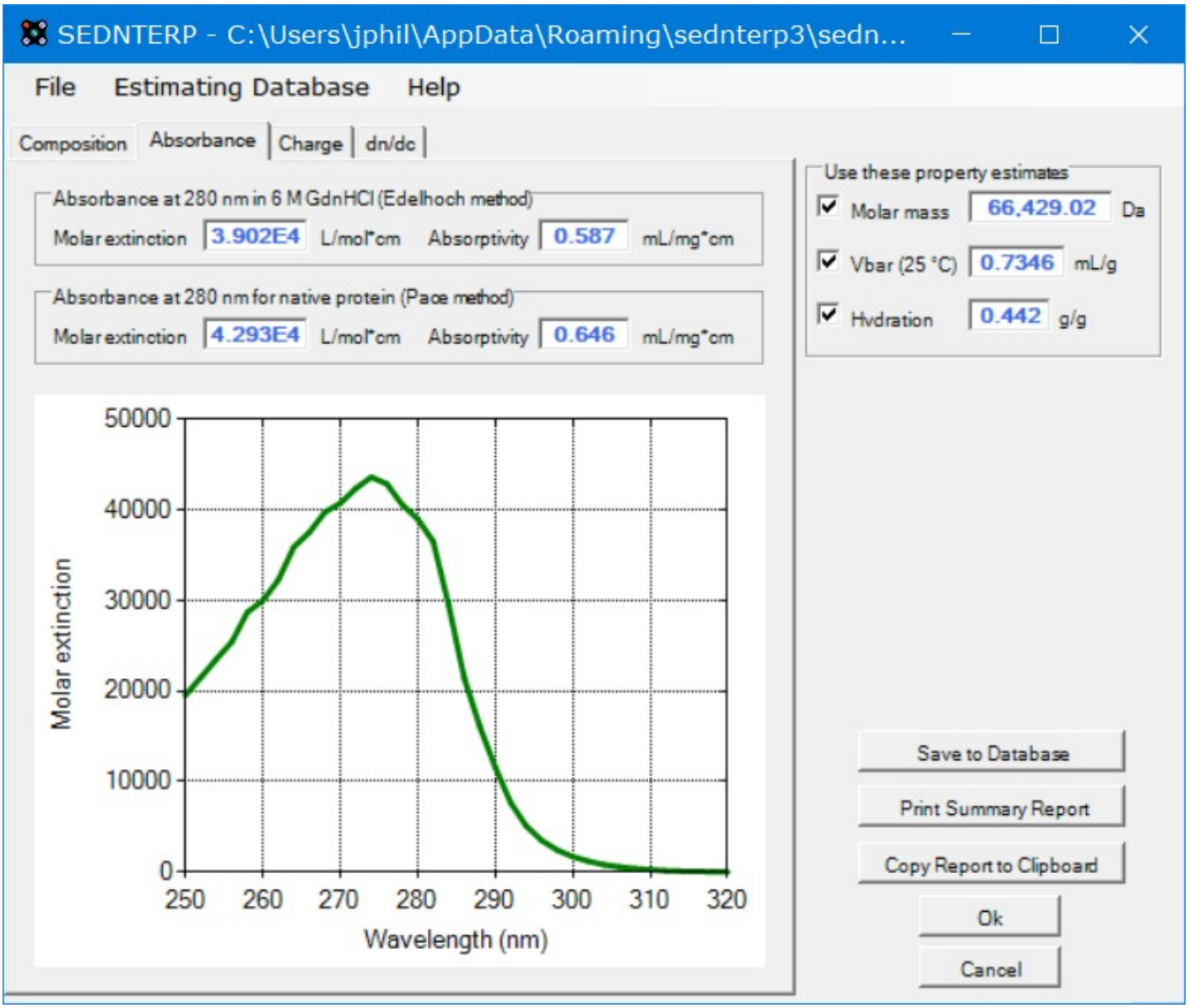
Since not all of the defined conjugates in the database have molar refractivity data (which is needed to calculate an accurate *dn/dc* value for the conjugated protein), the 'R' column of the conjugates table indicates which ones have such data. The checked rows have refractivity values, the un-checked ones with a salmon-colored background do not.

TIP: The conjugates list includes new items to help provide $\bar{\nu}$ and *dn/dc* estimates for glycoproteins and PEGylated proteins. The 'N-linked glycans' item represents 1 kDa of glycan with an average $\bar{\nu}$ value from the literature (taken from ref. 80). Thus if your glycoprotein on average contains 5.4 kDa of glycans, you would add 5.4 moles of this conjugate. Similarly the 'polyethylene glycol 10 kDa' item has $\bar{\nu}$ and refractivity values taken from ref. 81, and can be used to add the appropriate mass of PEG. The 'N-linked glycans' item adds hydration equivalent to that assumed for N-acetyl-glucosamine, while for 'polyethylene glycol 10 kDa' item the hydration data come from ref. 89.

The **Clear** button will remove all the moles entries from both the amino acid and conjugate tables.

The Absorbance tab

As illustrated below, clicking the **Absorbance tab** shows the molar extinction coefficient at 280 nm, and also the absorbance per mg/mL (absorptivity) at 280 nm (each computed by two different methods, detailed here). It also shows a graph of the estimated extinction spectrum.



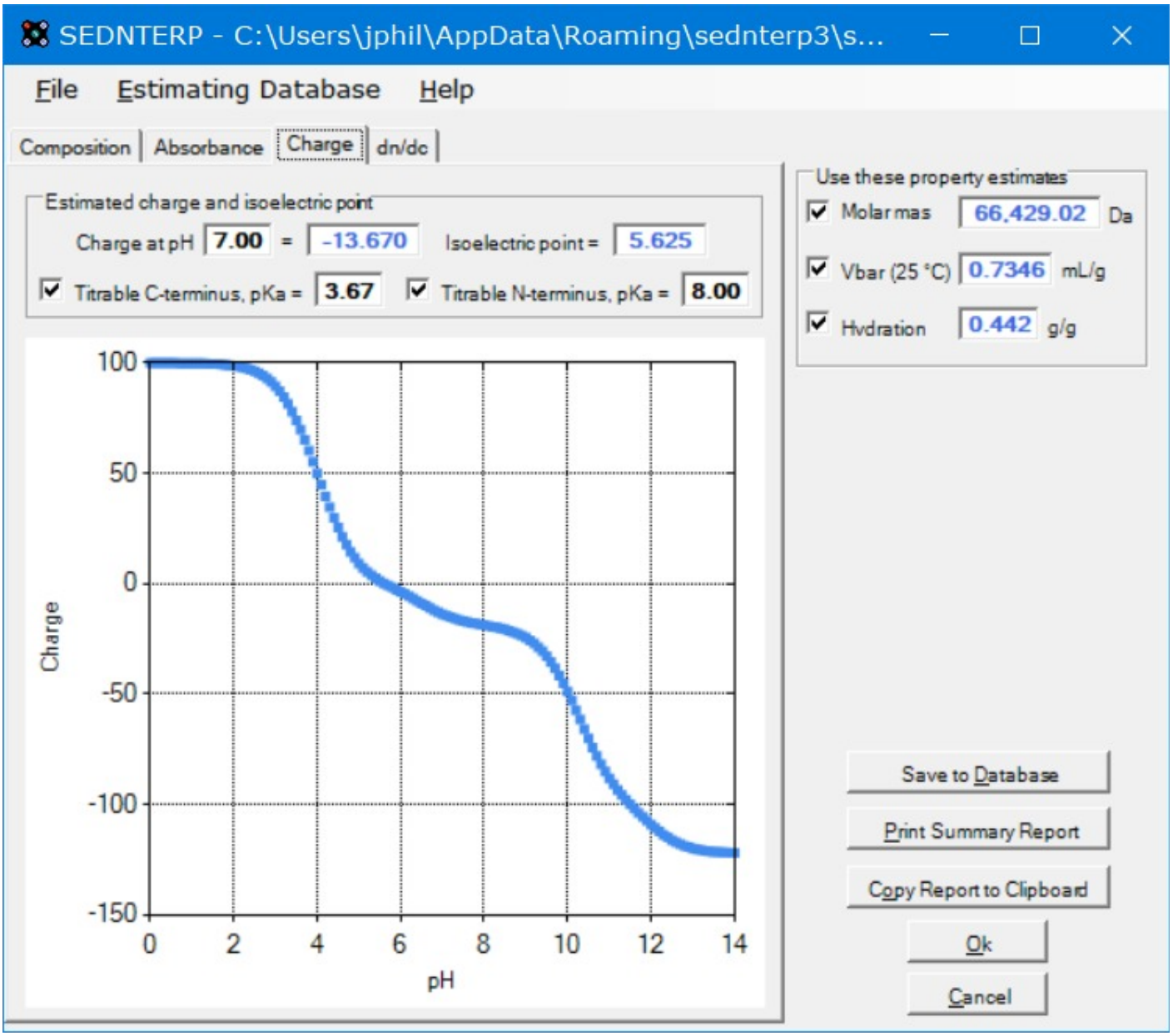
TIP: While these estimated extinction spectra may not be highly accurate for native proteins, experience shows that they often do predict whether the maximum absorbance for the native protein occurs at 280 nm (as in the BSA example shown above), and thus predict what wavelength should be used in AUC experiments to be at the top of the absorbance peak (and thus avoid potential artifacts due to wavelength shifts).

Note than none of these absorbance properties are used in other calculations, but this information may be useful in designing AUC experiments and determining the concentrations of your samples.

Note also that none of the defined conjugates currently have extinction spectral data in the database.

The Charge tab

As illustrated below, clicking the **Charge tab** shows the calculated charge at the experimental pH, the isoelectric point of the protein, and a graph of the estimated charge as a function of pH (titration curve).



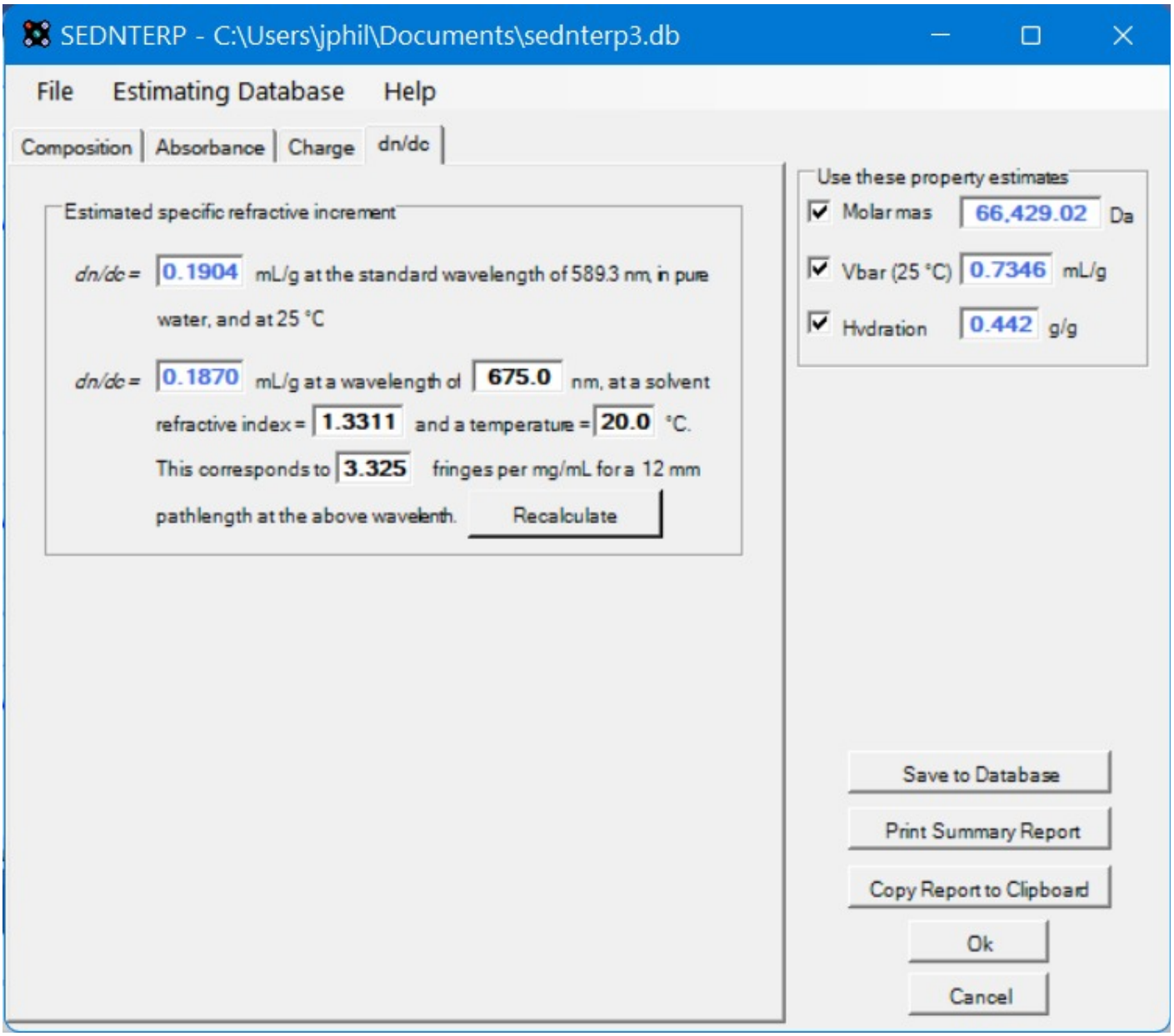
Changing the value in the **pH** text box will cause the program to calculate the charge at the new

pH. The user can also alter the default pK_a values for the amino terminus and carboxyl terminus, or make either terminus non-titrable by un-checking the corresponding check box.

Note than none of these charge properties are used in other SEDNTERP calculations, but this information may be useful in designing AUC or light scattering experiments. For example, avoiding a pH where the charge/mass ratio is quite high should help to minimize both thermodynamic and hydrodynamic solution non-ideality, and when you must work at a pH where the predicted charge is relatively high you may wish to increase the ionic strength to help reduce the non-ideality.

Caution: The actual net charge of a protein in solution may differ substantially from the theoretical values calculated by SEDNTERP. For example, it is known that monoclonal antibodies usually have a much lower net charge than the calculated value (see for example ref. 75). This difference from the theoretical charge is most likely due to ion binding by the protein.

The dn/dc tab



As illustrated above, clicking the ***dn/dc* tab** shows the estimated specific refractive increment. Two values are shown:

1. A 'standard' value at the standard wavelength of 589.26 nm, at 25 C, and in pure water.
2. An 'experimental' value at a chosen 'experimental' wavelength, in a buffer with a specified refractive index, and at a chosen temperature.

SEDNTERP then uses this 'experimental' *dn/dc* to calculate and display a factor used to convert AUC interferometer signals into weight concentration (the fringes per mg/mL for a cell with 12 mm pathlength). Any of these 'experimental' parameters can be changed by altering the values in

the wavelength, solvent refractive index, or temperature text boxes; the 'experimental' *dn/dc* value will be recalculated after you exit one of those text boxes with the ENTER or TAB keys, or when you click the **Recalculate button**. The initial values for the temperature and buffer refractive index are transferred from the current Buffer Properties tab (after adjusting the buffer RI value to the experimental wavelength).

TIP: The 'experimental' *dn/dc* wavelength is a default that can be altered by the user. The as-installed default experimental wavelength is 675 nm, which corresponds to a typical wavelength for the laser in the Rayleigh interferometer of an older XL-I AUC instrument. Thus this default wavelength may be appropriate for AUC users, but **probably is not correct for those users primarily interested in light scattering**. Furthermore, AUC users should be aware that these laser wavelengths vary somewhat from one instrument to another, and vary more significantly depending on the instrument model (XL-I *vs.* Optima) and XL-I manufacturing date. Thus AUC users should determine the actual wavelength for their interferometer by looking for a value given on the 'periscope' laser housing or by examining the 'header' (2nd line) of an interference scan file (the wavelength is the next-to-last entry on that line).

Here again none of the information on this tab is used in other SEDNTERP calculations, but these *dn/dc* values are quite important in interpreting static light scattering experiments, and should help AUC users to determine the true concentration of their samples.

The summary report

After the composition is complete, you can use the **Print Summary Report** button to print out a formatted summary (with tables) of various sample properties that can be computed from the composition, or the **Copy Report to Clipboard** button to place a copy of the same report onto the Windows clipboard (as formatted 'rich-text'). This report includes some things not shown on the Sample Composition form, such as the mean residue weight (needed for circular dichroism calculations) and the number and percentages of charged, acidic, basic, polar, and hydrophobic amino acids.

Returning values for use in other calculations

After the composition is entered, the molar mass, vbar, and/or hydration estimate may be returned to the main program. If for some reason, one or more of these values is **not** to be used, clearing

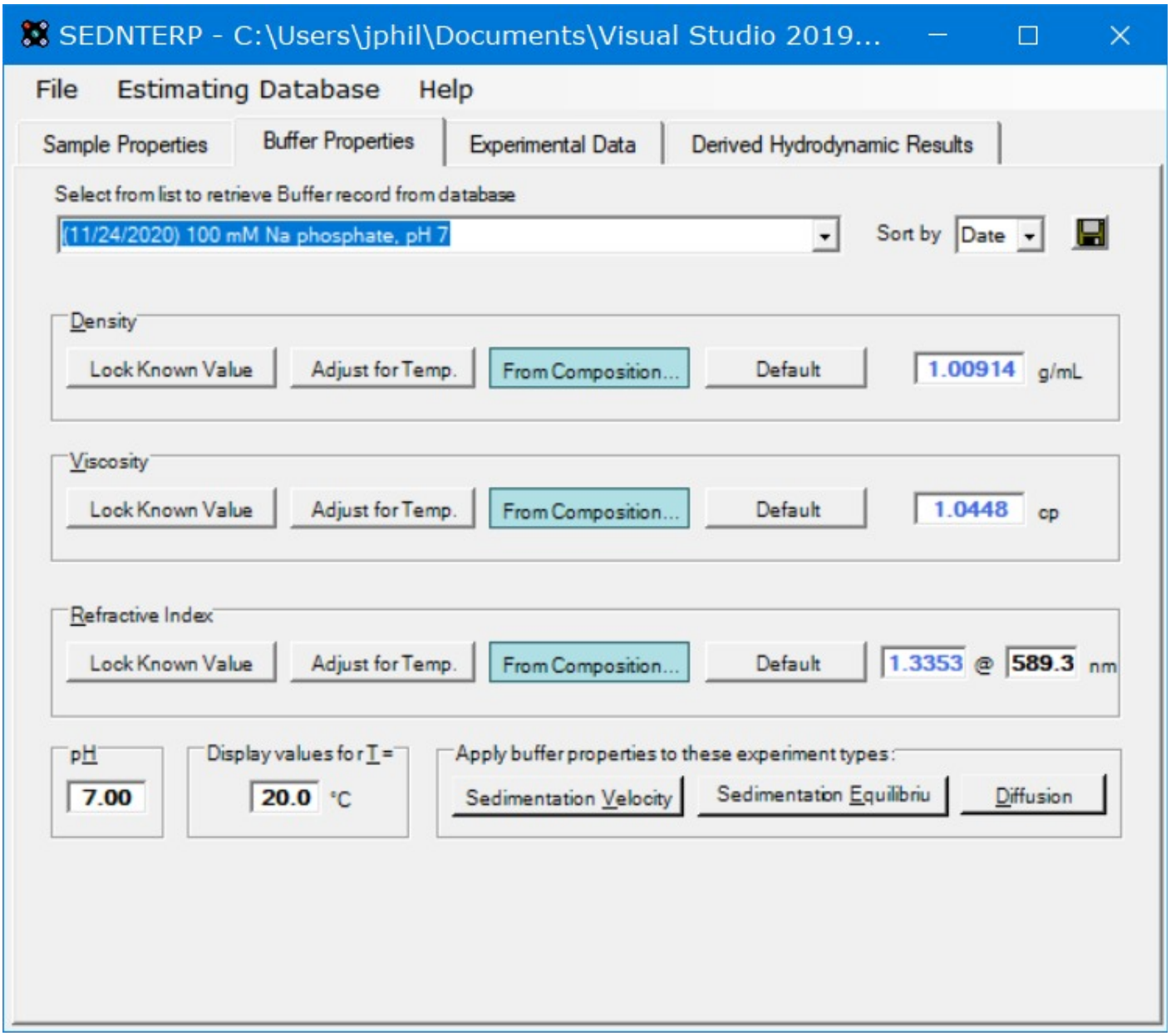
the check boxes will prevent SEDNTERP from transferring and using these values in further calculations. The values are returned when you use the **Ok** button. The **Cancel** button will return you to the main form, undoing all changes made while the Sample Composition form was active.

Saving this sample composition to the program database

The **Save to Database** button saves the changes and then jumps to The Saving Data in Database Form for recording this composition as a Sample record in the database. That is, the **Save to Sample Database** command is equivalent to pressing the **Ok** button on the Sample Composition form and then immediately selecting **Save sample Info to database** from the File Menu on the main form.

4.6 Buffer Properties major tab

The Buffer Properties major tab of the Main form displays information about properties of a particular buffer. The image below shows what this tab page looks like after the '100 mM Na phosphate, pH 7' record (included in the start-up database) has been loaded.



This tab page contains 3 main panels, each about a different sample property:

- density (ρ)
- viscosity (η)
- refractive index (*n*)

For each of those 3 properties, the source of the property value used by SEDNTERP in its calculations is indicated by which of the 4 buttons within that panel shows the light blue background. The meaning of these 4 buttons is the same for density, viscosity, and refractive index.

What do the sets of 4 buttons signify?

The **Lock Known Value** button indicates that user will enter a known ρ, η, or *n* value into the text box on the right, and also that these values should not be temperature corrected (or wavelength corrected for *n*) by SEDNTERP (this is indicated by the 'No T correction' message that appears in red below the corresponding text box).

The **Adjust for Temp.** button allows the user to enter a ρ, η, or *n* value at a particular temperature (and particular wavelength for *n*). Preferably that value is from an actual measurement you or someone else has made. The entered value will be then be corrected for any temperature (or wavelength) difference between the entered value and the experiment. When this button is pressed it brings up a small dialog box for entering the known value and its measurement temperature (and wavelength for *n*).

The **From Composition...** button indicates that ρ, η, or *n* has been computed from the buffer's composition (the types of solutes and their concentrations). When the button is clicked the 4 main tabs are temporarily replaced by the Buffer Composition form so the user can enter the composition information.

The **Default** button sets ρ, η, or *n* to the current default value (which are the density, viscosity, and refractive index for water when the program is installed, but which can be changed to whatever the user prefers).

The 'Apply to these experiment types' buttons

SEDNTERP internally maintains separate sets of buffer properties for sedimentation velocity, sedimentation equilibrium, and diffusion experiments, as well as a 'current' buffer (the one whose properties are currently being displayed on the Buffer Properties major tab. At the bottom right are three toggle buttons which when turned on (clicked so they display a light blue background) will transfer the density and/or viscosity of the 'current' buffer to one of the 3 experiment types. While a button is on, any further revisions to the buffer properties will transfer to the experiment page (note that only buffer density is relevant to equilibrium experiments, and only buffer viscosity is relevant to diffusion experiments). Toggling a button off will stop automatic transfer of changes in buffer properties to that experiment type.

The pH and Temperature buttons

At the bottom left are two text boxes. The first text box allows the user to specify the pH for the current buffer. Note that this value is merely descriptive---it is not actually used in any of the calculations. Also note that setting the pH to a value <= 4 does not automatically select the 'Calculate hydration for pH <= 4' option on the Sample Composition form.

The other text box allows the user to change the temperature for which the density, viscosity, and refractive index are being displayed.

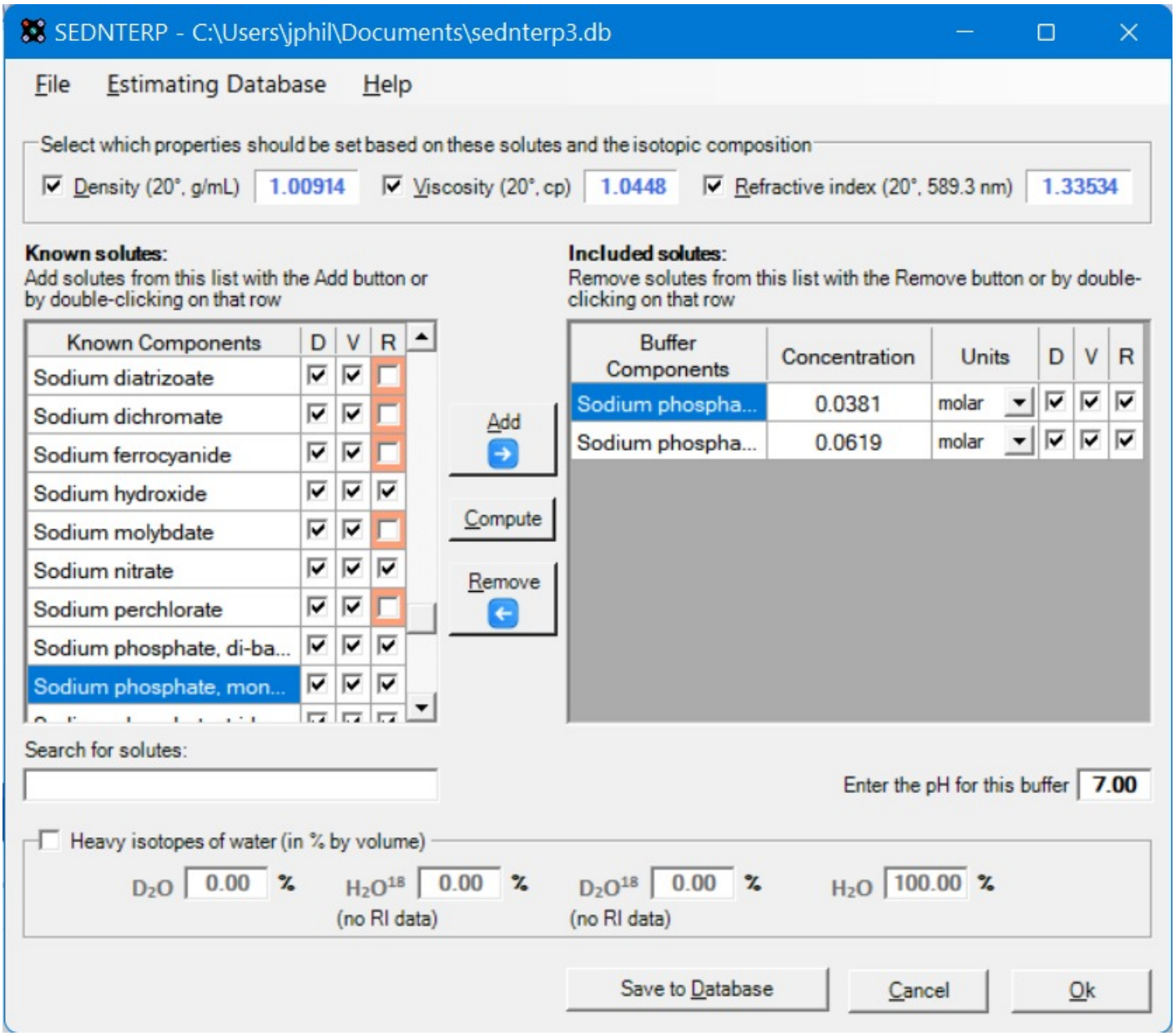
Retrieving or saving a Buffer record from/to the database

A previously saved Buffer record can be retrieved from the database when chosen by name from a drop-down list box located at the upper left in this section. Whether that list is sorted by name, or by the date the Buffer record was last modified, is controlled by the **Sort by** drop-down list to its right. The **Save** button at the upper right (with a disk icon) will initiate saving a new Buffer record (or over-writing the current one) by bringing up the Saving Data in Databases form.

4.7 Buffer Composition form

The purpose of the Buffer Components form is to allow the user to specify the components (ingredients) of a buffer and then to have SEDNTERP interpolate a buffer density, viscosity, and/or refractive index for that buffer. Standard tables of single component buffer data have been fitted by polynomials, and the coefficients of those polynomials are stored in the physical constants database. SEDNTERP computes density, viscosity, or refractive index increments for each component at the user entered concentration, and then adds these increments together to compute a final buffer density, viscosity, or refractive index. See the density interpolation, viscosity interpolation, and RI interpolation pages for more complete information on the formulae involved.

The image below shows what this form looks like when the density is calculated for the '100 mM sodium phosphate, pH 7.0' buffer stored in the startup database installed with the program.



What information do you need about your buffer?

It is important to realize that SEDNTERP is **not** a tool that calculates "recipes" for buffers. Rather, it assumes that you know the full composition information for your buffer, including (for best accuracy) the concentration of both the acidic and basic forms of your buffering agent (for example, the mono- and di-basic forms of sodium phosphate for the buffer shown above). Therefore if you do not have that information at hand, prior to creating your buffer in SEDNTERP you may want to use a "buffer recipe" calculator tool that performs such calculations. One recommended web-based tool is Bob Beynon's 'Calculator for pH buffers'

at https://www.liverpool.ac.uk/pfg/Research/Tools/BuffferCalc/Buffer.html. Another widely-used tool based on Excel spreadsheets is CurTiPot, available at http://www.iq.usp.br/gutz/Curtipot_.html.

How to add components

Choose buffer components (solutes) from the list provided on the left-hand side by either double-clicking on the solute name, or first selecting from the list and then clicking the **Add button**. Next, enter the concentration of the component, and the units for that concentration, on the new row that has been added to the 'Included solutes' table on the right-hand side. For all solutes the concentration units may be molar, molal, mg/mL, g/mL, or % by weight. For solutes that are liquids, units of % by volume are also allowed. Note that internally the program always converts your concentration entry into molar units, and that not all components need to use the same concentration units. The equations and computation methods used to convert entered concentrations to molar units are described at the end of this topic.

TIP: When you have entered a concentration in units other than molar, you can see the equivalent molar concentration computed by the program by hovering your mouse over the table cell where you entered the concentration (the molar concentration will then show as a ToolTip).

Components entered mistakenly may be removed with the **Remove button**.

If the 'D', 'V', or 'R' column checkbox next to the solute name in the 'Known solutes' table is unchecked and has a salmon-colored background, that means its density, viscosity, or RI data (respectively) is not available, or has not yet been fitted to the polynomial function SEDNTERP uses to interpolate these values. If you have experimental data available, new components may be added to the database of interpolating polynomials. If a component is used for which some data is missing, the background color for the text box at the top of the page where the interpolated property value is displayed will change to salmon to alert the user.

Note: The most frequently asked question about SEDNTERP is "Why doesn't SEDNTERP have data for [X, some buffering agent or salt]?". The answer is always "sorry, if it isn't in the database, that means we aren't aware of any density or viscosity data for that compound". Remember that it is usually reasonable to neglect contributions of compounds present only at concentrations of a few mM. Remember too that it is not too difficult to measure buffer densities with sufficient precision for this purpose (see the Tip box in Interpolating a Buffer Density). On the other hand, measuring the

viscosity with an accuracy comparable to the ~0.2% precision of sedimentation coefficients is actually beyond the capability of most viscometers.

The maximum molar concentration covered by the polynomial fit to the original experimental data table varies from one solute to another, and might be different for density, viscosity, and refractive index. If a concentration higher than the available interpolation maximum is entered, SEDNTERP will allow the calculation, but will alert the user that it may be inaccurate by highlighting the calculated value in yellow, and also changing the background color to yellow for the D, V, or R checkbox on the right side of that solute entry in the **Added Solutes table**.

Typing in the Search box at the bottom of the 'Known solutes' table initiates a alphabetical search for whatever is typed in the search box. This enables the user to quickly find a particular solute in the long list.

Specifying heavy isotopes of water

The unchecked box next to the 'Heavy isotopes of water' panel name at the bottom of the form indicates that this buffer does not contain any heavy isotopes of water. If the buffer does have heavy water in it, the density calculation can account for the density of heavy water components. Checking this box will activate the text boxes where the percentages of D_2O, H_2O^{18}, and D_2O^{18} can be entered (the percentage of normal water is automatically computed).

CAUTION: Although SEDNTERP can calculate densities for buffers containing H_2O^{18} or D_2O^{18}, it does not have viscosity or refractive index data for those isotopes (as is noted below those text boxes on the form).

pH

A text box below the 'Includes solutes' table allows the user to specify the pH for this buffer (this value will be tranferred to the Buffer Properties major tab). Note that this value is merely descriptive---it is not actually used in any of the calculations. Also note that setting the pH to a value <= 4 does not automatically select the 'Calculate hydration for pH <= 4' option on the Sample Composition form.

Returning interpolated values to the Main form

The interpolated values for the density, viscosity, and refractive index, including the effects of any heavy isotopes of water, are displayed along the top of the form. Those values shown are for

20 ^{O}C.

The check box to the left of each interpolated value determines whether that property will be assigned to the current buffer shown on the Buffer Properties major tab when the **Ok button** is clicked to close the Buffer Composition form.

Clicking the **Cancel button** will return to the Buffer Properties major tab without changing its state.

Other controls

A previously saved Buffer record can be retrieved from the database when chosen by name from a drop-down list box located at the upper left in this section. Whether that list is sorted by name, or by the recorded experiment date, is controlled by the **Sort by** drop-down list to its right. The **Save** button at the upper right (with a disk icon) will initiate saving a new Buffer record (or over-writing the current one) by bringing up the Saving Data in Databases form.

Allowing for buffer solute concentration entry in units other than molar

A new feature in version 3 of SEDNTERP allows the user to enter the solute concentration in units of molal, mg/mL, g/mL, or % by weight. For solutes that are liquids, units of % by volume are also allowed. Conversion from mg/mL, g/mL, or % by volume to molar units is straightforward, and was made possible simply by adding the solute molar mass to the program database, and for the liquid solutes also adding their density.

Converting from molal or % by weight units to molar, however, is much more complicated. For a buffer containing n components at molal concentrations b_i, the formula to convert to the equivalent set of molar concentrations C_i is:

Equation 28:

$$C_i = \frac{\rho b_i}{1 + \sum_{j=1}^{n} b_j M_j}$$

where ρ is the density of that buffer solution. Similarly, for solutes at weight percent w_i the conversion to molar is given by:

Equation 29:

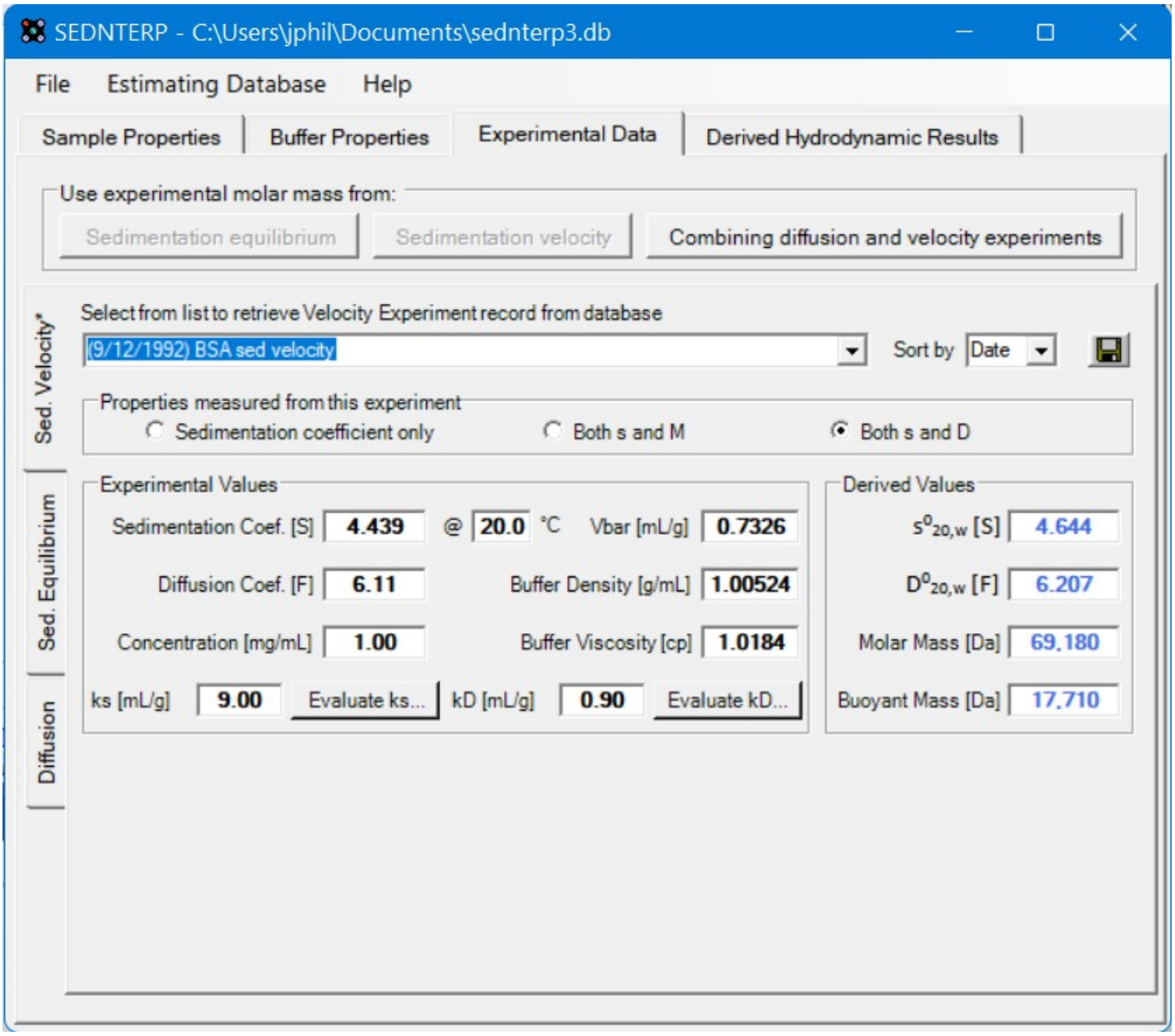
$$C_i = 10 \times \frac{\rho w_i}{M_i}$$

These two equations show that the solution density must be known to make those conversions, but that density is exactly what we are trying to compute. This means that for these two cases the molar concentrations must be determined iteratively: an initial guess is made for the molar concentration, the total buffer density is updated based on that value, a revised molar concentration is calculated based on that updated density, and then the process repeats until the last iteration changes the molar concentration by less than 0.001%. Since every component affects the density and hence the molar concentration of any solute entered as molal or wt%, that means that those molar concentrations must be recalculated via iteration each time an additional component is added to the buffer, or when the concentration of an existing component is changed. Furthermore, when more than one component is entered as molal or wt%, two levels of iteration are required: an outer iteration loop where the total buffer density is updated based on current estimates for the molar concentrations of all components entered in molal or % by weight units, and an inner iteration loop where the molar concentration of each such component is recomputed based on the current total buffer density.

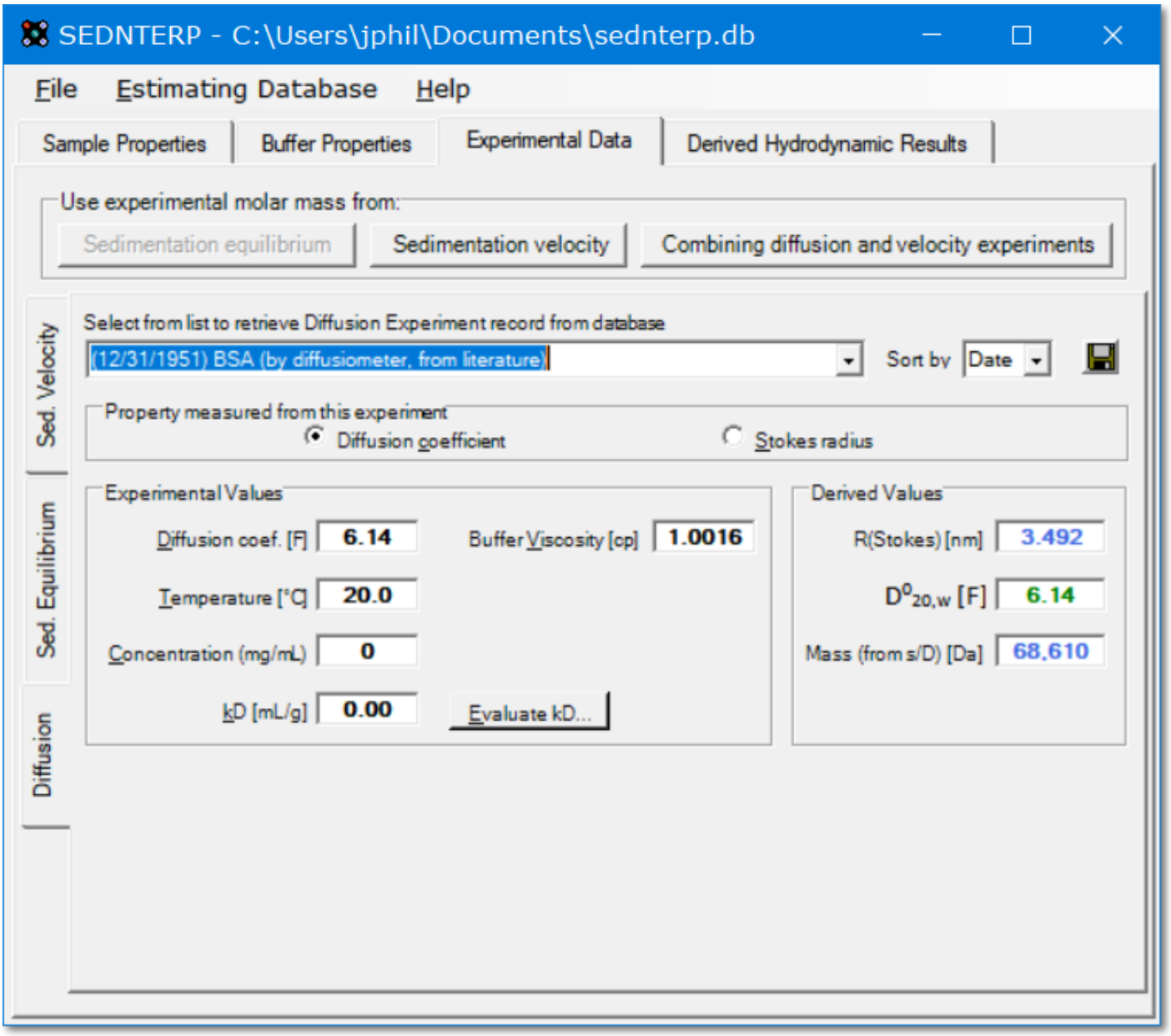
4.8 Experimental Data major tab

The Experimental Data major tab of the Main form displays information about the results from, and conditions for, sedimentation velocity, sedimentation equilibrium, and/or diffusion experiments (data for all three types can be present simultaneously). Note also that each of the three experiment types can use a different buffer, but all experiment types loaded simultaneously must share a common set of Sample Properties (*i.e.* in the database, all the experiment records would normally be linked to the same sample record).

The image below shows what this tab page looks like after the '(9/12/1992) BSA sed velocity' record (included in the start-up database) has been loaded.



That velocity experiment record is also linked to a diffusion experiment record (data from the literature), which gets automatically loaded along with this BSA velocity record. That diffusion experiment information can be displayed by clicking on the 'Diffusion' tab along the left, as shown below:



The details about the inputs and interface controls for each type of experiment are discussed elsewhere (see Sedimentation Velocity minor tab, Sedimentation Equilibrium minor tab, and Diffusion minor tab). The remainder of this topic will discuss some general features of experiments and experiment types.

About the linkage between experiment type display and the Buffer Properties tab

Note that the displayed buffer viscosity differs between the diffusion and velocity experiments (because they were in different buffers). SEDNTERP internally maintains a different set of buffer properties for each type of experiment. When multiple experiment types are loaded, and when

those experiments are linked to a buffer record (as is normally the case), which buffer record gets displayed when you go to the Buffer Properties major tab after displaying the Experimental Data tab is determined by which type of experiment was last displayed. Thus, in this example if we last displayed the velocity experiment, then its buffer (Dulbecco's PBS) will be displayed on the Buffer Properties tab, whereas if we last displayed the diffusion experiment, we would see its buffer on the Buffer Properties tab (its 'buffer' is actually water because the literature reported values were extrapolated to water).

Choosing which experimental molar mass should be used in further calculations

When results from multiple experiment types are present, it may be possible to derive an experimental molar mass in more than one way. Therefore a set of 3 buttons is provided near the top of the Experimental Data tab page, within the 'Use experimental molar mass from' panel, to control which (if any) experimental molar mass should be used. Which molar mass is currently being used is indicated by a light blue background color for the corresponding button.

The **Sedimentation Equilibrium** button is active whenever sedimentation equilibrium data is available.

The **Sedimentation Velocity** button is active when the sedimentation velocity experiment returned values for both *s* and *D*, or a value for *M* as well as one for *s*.

The **Combining diffusion and velocity experiments** button chooses a method for deriving an experimental molar mass that is new to SEDNTERP 3. This approach combines a sedimentation coefficient with a diffusion coefficient from a separate diffusion experiment, and calculates *M* from the Svedberg equation:

$$M = \frac{sRT}{D\left(1-\bar{v}\rho\right)}$$

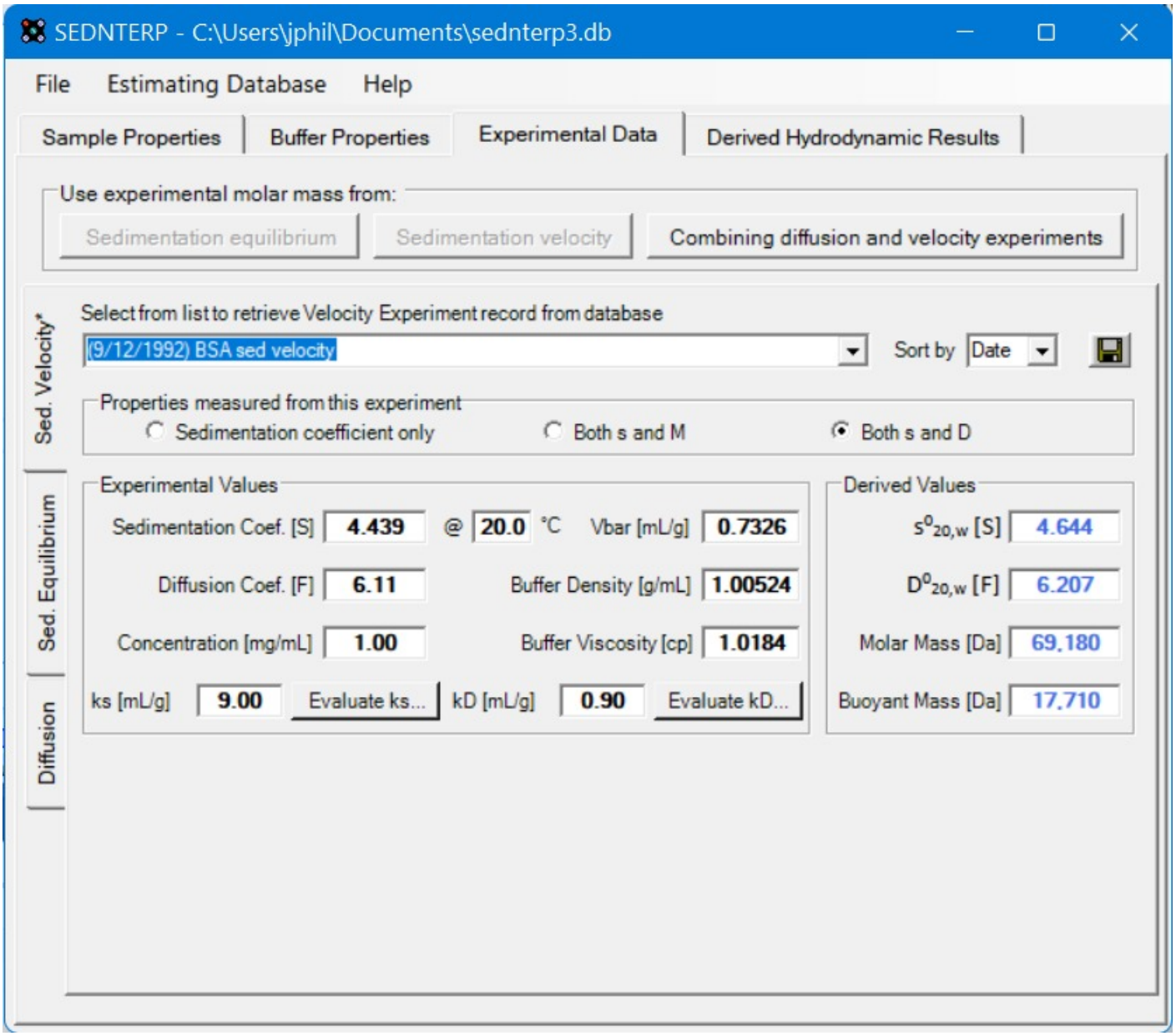
Note: One advantage of this combined diffusion/sedimentation approach is that it can (at least in principle) determine the average molar mass for samples that are intrinsically heterogeneous in molar mass (for example a polymer), and where the measured sedimentation and diffusion coefficients each represent an average over the multiple species that are present. In theory, to obtain a true weight-average molar mass requires combining a true weight-average sedimentation coefficient with a true *z*-average diffusion coefficient (ref. 91), which is what dynamic light scattering would determine (assuming all the species have the same *dn/dc*).

When two sedimentation velocity and diffusion experiments are being combined, SEDNTERP

provides a visual reminder that they are being linked together by displaying a link icon  on the sedimentation velocity and diffusion tabs on the left side.

4.9 Sedimentation Velocity minor tab

The Sedimentation Velocity minor tab of the Experimental Data major tab displays information about the results from, and conditions for, sedimentation velocity experiments. The image below shows what this tab page looks like after the '(9/12/1992) BSA sed velocity' record (included in the start-up database) has been loaded.



First, note that the buttons within the **Use experimental molar mass from:** panel above the Sed. Velocity tab are discussed in the Experimental Data major tab topic.

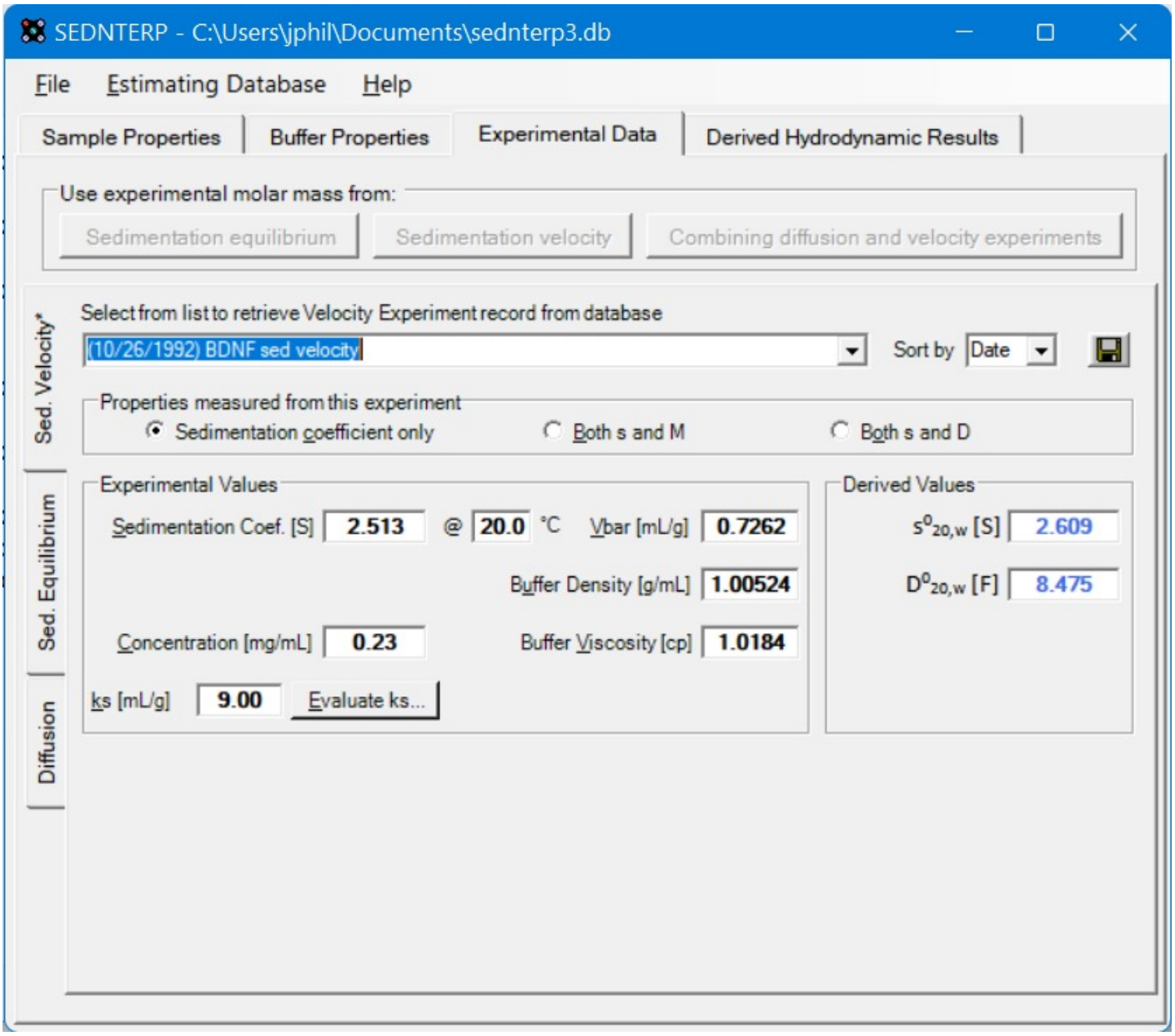
The 'Properties measured from this experiment' radio buttons

This group of three radio buttons allows to user to specify which molecular parameters were returned from the data analysis of this sedimentation velocity experiment. Those options in turn govern which inputs and buttons appear within the **Experimental Values** panel, and also which outputs appear in the **Derived Values** panel (both below).

The three options are:

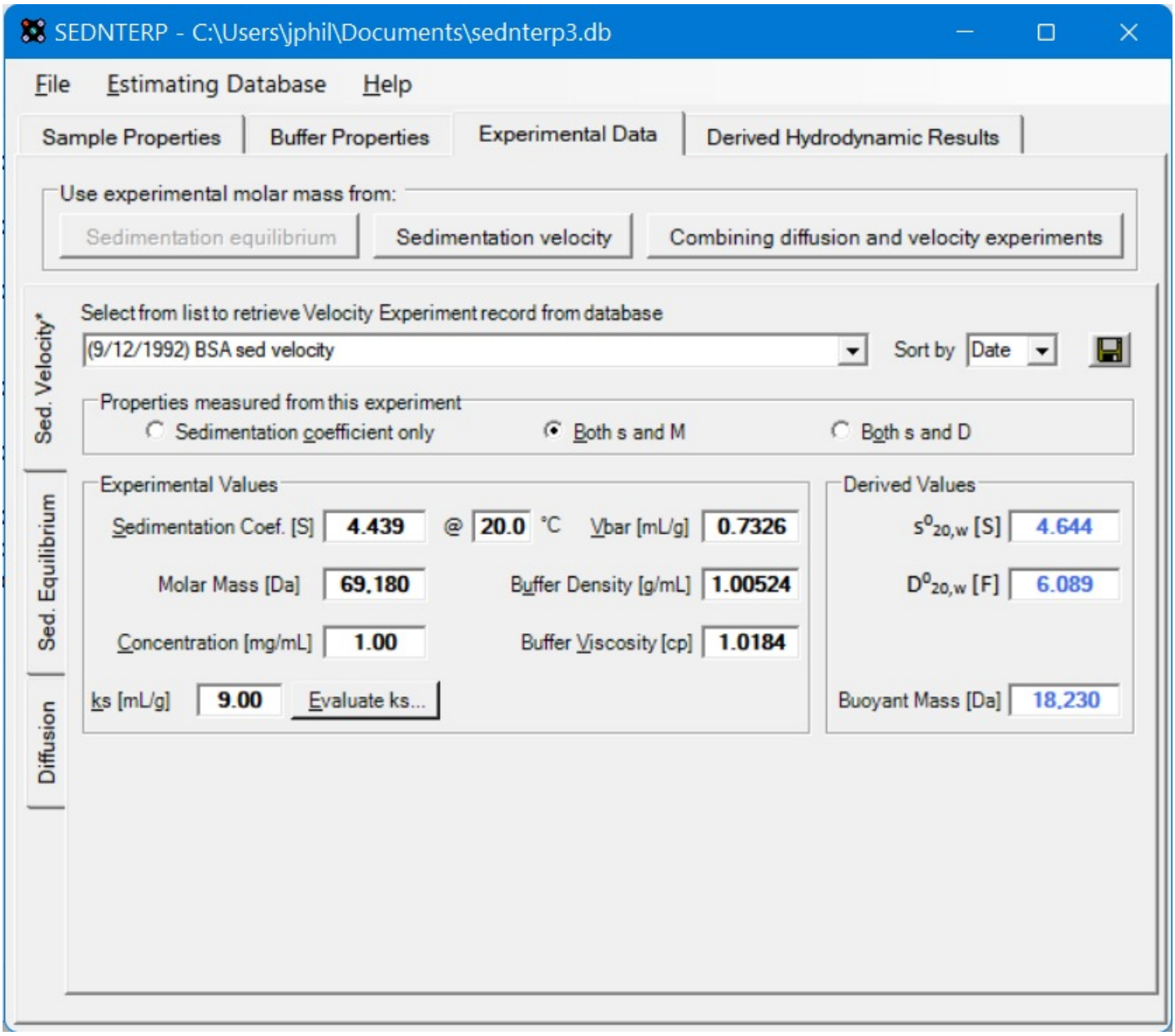
1. **Sedimentation coefficient only**, which obviously means that the velocity experiment yielded only an *s** value
2. **Both *s* and *M***, which means the velocity data analysis returned an estimate for the molar mass as well as an *s** value
3. **Both *s* and *D***, which means the velocity data analysis returned an estimate for the diffusion coefficient as well as an *s** value

The image below shows what this tab looks like for option 1:



Note that the $D^0_{20,w}$ value displayed is calculated from $s^0_{20,w}$ and the known molar mass (from composition in this case).

The image below shows what this tab looks like for option 2:



In this case the $D^0_{20,w}$ value displayed is calculated from $s^0_{20,w}$ and the experimental molar mass (taken from this experiment). The buoyant mass is also calculated and displayed.

For option 3 (image at top), the s^* and D^* inputs are directly converted to $s^0_{20,w}$ and $D^0_{20,w}$ values,

and the ratio of those values is used to calculate an experimental molar mass and buoyant molar mass.

Overriding a computed $s^0_{20,w}$ or $D^0_{20,w}$ value

If you know the correct $s^0_{20,w}$ or $D^0_{20,w}$ value from another source, it is possible to override the values that SEDNTERP computes and directly enter the known value into the text boxes within the 'Derived Values' panel. Such 'user override' values are displayed in green text. (Note that you cannot override the $D^0_{20,w}$ value that is computed for a 'sedimentation coefficient only' experiment since that value is not computed from an experimental diffusion coefficient.)

Experimental condition inputs

Input text boxes are provided for the experimental temperature, as well as the $\bar{\nu}$, buffer density, and buffer viscosity at the experimental temperature. Those values are normally transferred from the current ones on the Sample and Buffer Properties tabs, but can be altered by the user.

Concentration, k_s, and k_D inputs

The sample concentration input is used together with the values of k_s (and k_D also when D^* was measured) to extrapolate to zero concentration and calculate the displayed $s^0_{20,w}$ and $D^0_{20,w}$ values. If no concentration is entered (or if it is set as zero) then no concentration extrapolation is made, and then $s_{20,w}$ and $D_{20,w}$ values are displayed instead.

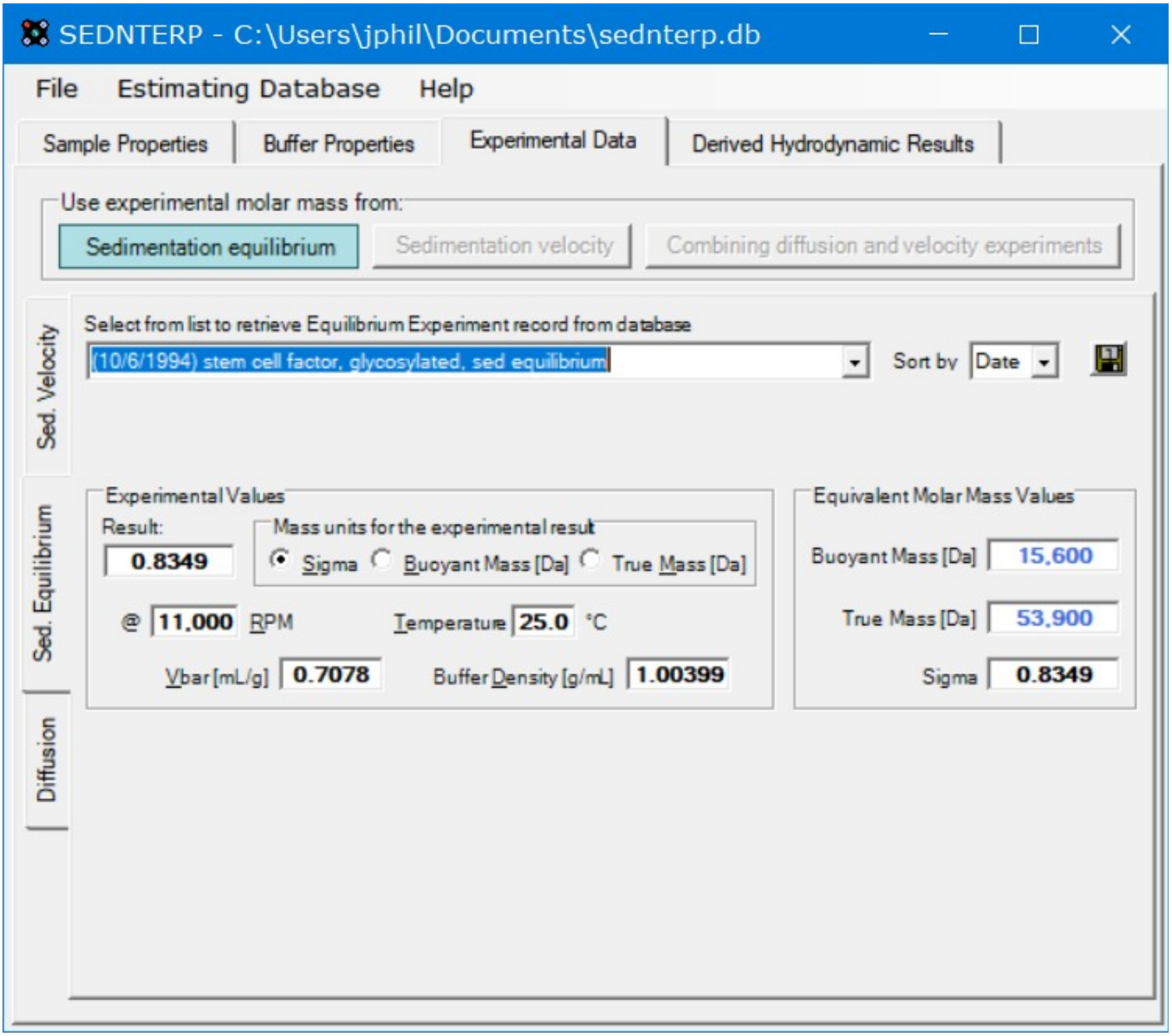
The k_s and k_D values are normally transferred from the Sample Properties tab, but the user can enter a different value, or click the **Evaluate ks** or **Evaluate kD** buttons to bring up a form for entering s or D values at multiple concentrations in order to experimentally determine k_s or k_D. From that form the user can optionally also choose to use the extrapolation of the entered dilution series to zero concentration as the final $s^0_{20,w}$ or $D^0_{20,w}$ value.

Retrieving or saving a Sedimentation Velocity Experiment record from/to the database

A previously saved Sedimentation Velocity Experiment record can be retrieved from the database when chosen by name from a drop-down list box located at the upper left in this section. Whether that list is sorted by name, or by the recorded experiment date, is controlled by the **Sort by** drop-down list to its right. The **Save** button at the upper right (with a disk icon) will initiate saving a new Sedimentation Velocity Experiment record (or over-writing the current one) by bringing up the Saving Data in Databases form.

4.10 Sedimentation Equilibrium minor tab

The Sedimentation Equilibrium minor tab of the Experimental Data major tab displays information about the results from, and conditions for, sedimentation equilibrium experiments. The image below shows what this tab page looks like after the '(10/6/1994) stem cell factor, glycosylated, sed equilibrium' record (included in the start-up database) has been loaded.



Selecting the mass units

Different sedimentation equilibrium data analysis packages may return the results in different units, and therefore SEDNTERP now allows for entering the results as true molar mass, buoyant molar mass, or σ (reduced apparent molar mass). Which units will be used is governed by 3 radio buttons within the 'Mass units for the experimental result' panel. Note that when using σ units, calculating the buoyant and true molar mass requires that the rotor speed is also entered.

No matter which units are chosen, the equivalent molar mass values for all three types of units are displayed within the 'Equivalent Molar Mass Values' panel on the right side (but again calculating σ requires a rotor speed).

Experimental condition inputs

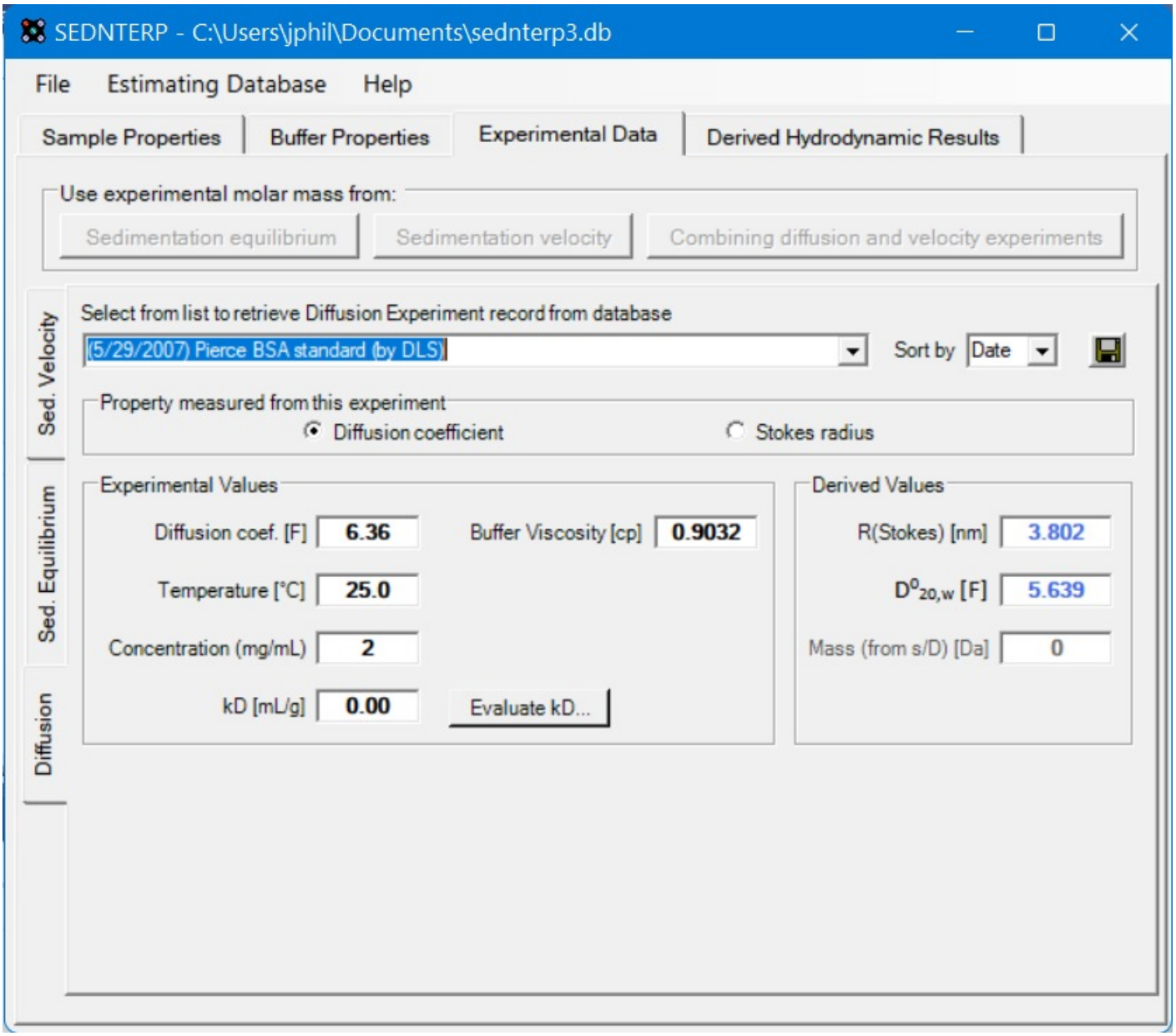
Input text boxes are provided for the experimental temperature, and the $\bar{\nu}$ and buffer density at that experimental temperature. Those values are normally transferred from the current ones on the Sample and Buffer Properties tabs, but can be altered by the user.

Retrieving or saving a Sedimentation Equilibrium Experiment record from/to the database

A previously saved Sedimentation Equilibrium Experiment record can be retrieved from the database when chosen by name from a drop-down list box located at the upper left in this section. Whether that list is sorted by name, or by the recorded experiment date, is controlled by the **Sort by** drop-down list to its right. The **Save** button at the upper right (with a disk icon) will initiate saving a new Sedimentation Equilibrium Experiment record (or over-writing the current one) by bringing up the Saving Data in Databases form.

4.11 Diffusion minor tab

The Diffusion minor tab of the Experimental Data major tab displays information about the results from, and conditions for, diffusion experiments. The image below shows what this tab page looks like after the '(5/29/2007) Pierce BSA standard (by DLS)' record (included in the start-up database) has been loaded.



The 'Property measured from this experiment' radio buttons

This group of two radio buttons allows to user to specify whether the diffusion experiment data analysis software returned a diffusion coefficient or instead returned a Stokes radius. That choice in turn governs which inputs and buttons appear within the **Experimental Values** panel, and which outputs appear in the **Derived Values** panel.

If the experiment returned D^*, then SEDNTERP calculates and displays the equivalent Stokes radius (R_s). If the experiment returned R_s , then SEDNTERP calculates and displays D^*. In either case $D_{20,w}$

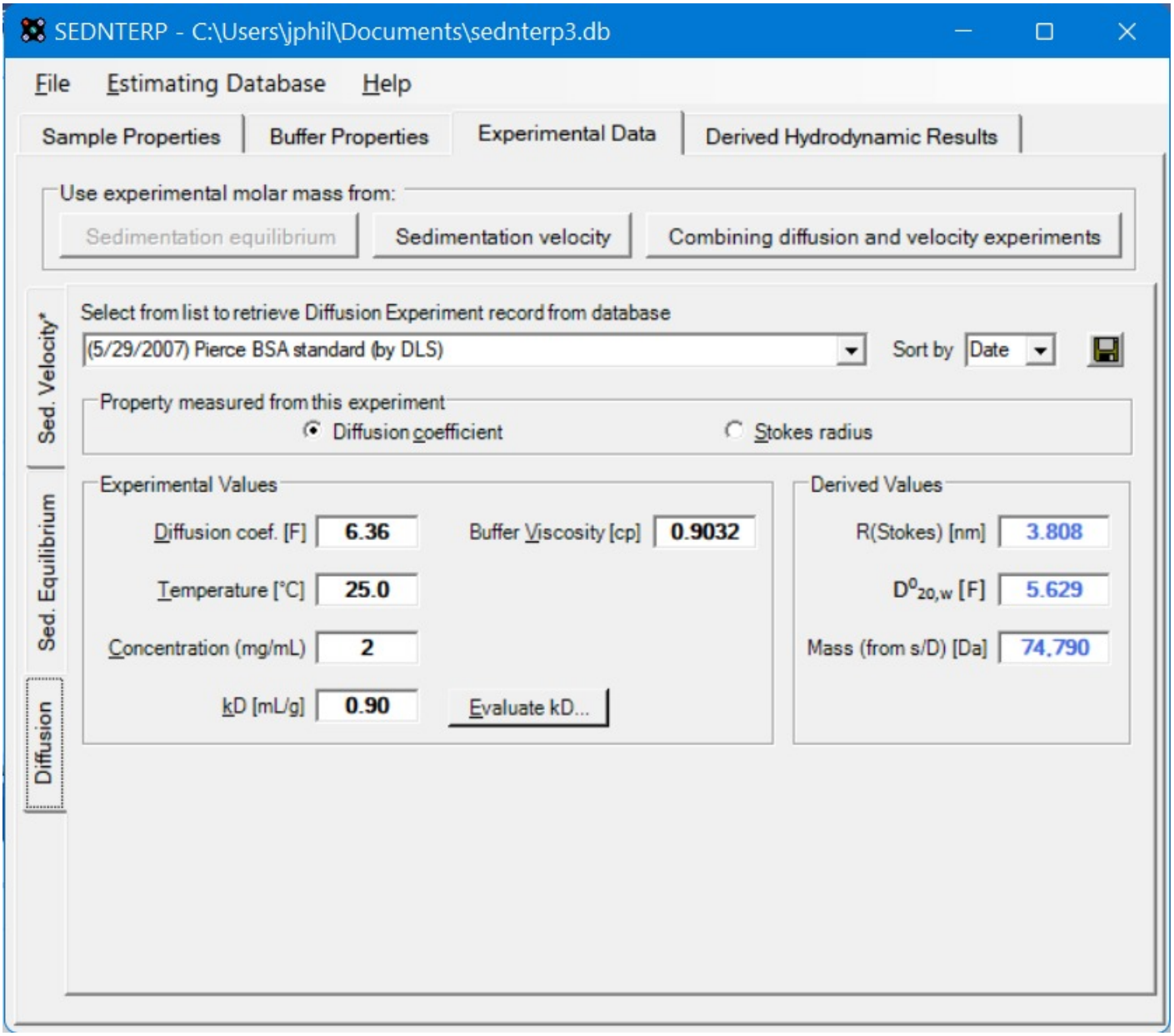
or $D^0_{20,w}$ is calculated and displayed.

The 'Mass (from s/D)' derived value

This derived value can only be calculated when a sedimentation coefficient for the same sample is currently present on the Sedimentation Velocity minor tab. When both *s* and *D* are known, the molar mass *M* can be calculated using the Svedberg equation:

$$M = \frac{sRT}{D\left(1-\bar{v}\rho\right)}$$

The image below shows this tab after a sedimentation coefficient for BSA has been loaded from the '(9/12/1992) BSA sed velocity' record (included in the start-up database), followed by loading a diffusion coefficient for BSA from the '(5/29/2007) Pierce BSA standard (by DLS)' record.



Note: In the example shown above, the calculated molar mass of 74.79 kDa is definitely higher than expected for BSA monomer. That error is most likely due to the presence of a substantial fraction of dimer (as well as some larger aggregates) in the Pierce BSA concentration standard. When using dynamic light scattering those aggregates cannot be resolved from the monomer, but the lower diffusion coefficients for the aggregates will still result in the DLS instrument reporting a lower (average) diffusion coefficient. (The poor resolution of DLS is a distinct drawback.)

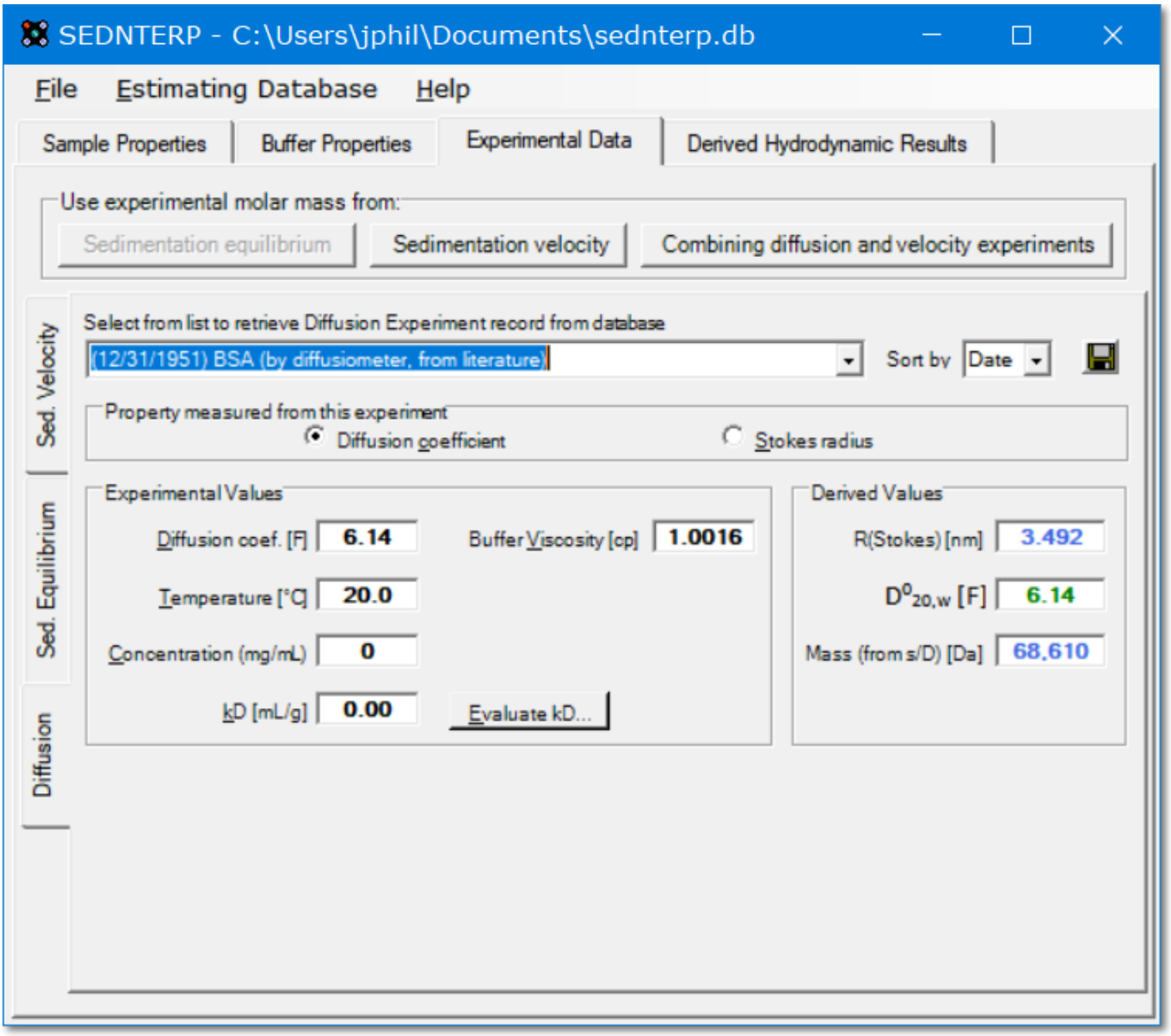
Nonetheless combining the sedimentation and diffusion experiments does correctly tell us that the predominant species is a monomer, and this derived mass is completely independent of the shape of the molecule (unlike the mass estimates usually provided by DLS software).

This molar mass calculated from *s/D* can be assigned as the molar mass value to be used in other calculations by clicking the **Combining diffusion and velocity experiments** button near the top of the Diffusion minor tab.

Note: One advantage of this combined diffusion/sedimentation approach is that it can (at least in principle) determine the average molar mass for samples that are intrinsically heterogeneous in molar mass (for example a polymer), and where the measured sedimentation and diffusion coefficients each represent an average over the multiple species that are present. In theory, to obtain a true weight-average molar mass requires combining a true weight-average sedimentation coefficient with a true *z*-average diffusion coefficient (ref. 91), which is what dynamic light scattering would determine (assuming all the species have the same *dn/dc*).

Overriding a computed $D^0_{20,w}$ value

If you know the correct $D^0_{20,w}$ value from another source, it is possible to override the value that SEDNTERP computes and directly enter the known value into the text box within the 'Derived Values' panel. Such 'user override' values are displayed in green text. An example of such an override is the '(12/31/1951) BSA (by diffusiometer, from the literature)' record (included in the start-up database), which is illustrated below:



Experimental condition inputs

Input text boxes are provided for the experimental temperature and the buffer viscosity at the experimental temperature. Those values are normally transferred from the current ones on the Sample and Buffer Properties tabs, but can be altered by the user.

Concentration and k_D inputs

The sample concentration input is used together with the value of k_D to extrapolate to zero

concentration and calculate the displayed $D^0_{20,w}$ value. If no concentration is entered (or if it is set as zero) then no concentration extrapolation is made, and then a $D_{20,w}$ value is displayed instead.

The k_D value is normally transferred from the Sample Properties tab, but the user can enter a different value, or click the **Evaluate kD** button to bring up a form for entering D values at multiple concentrations in order to experimentally determine k_D. From that form the user can optionally also choose to use the extrapolation of the entered dilution series to zero concentration as the final $D^0_{20,w}$ value.

Retrieving or saving a Diffusion Experiment record from/to the database

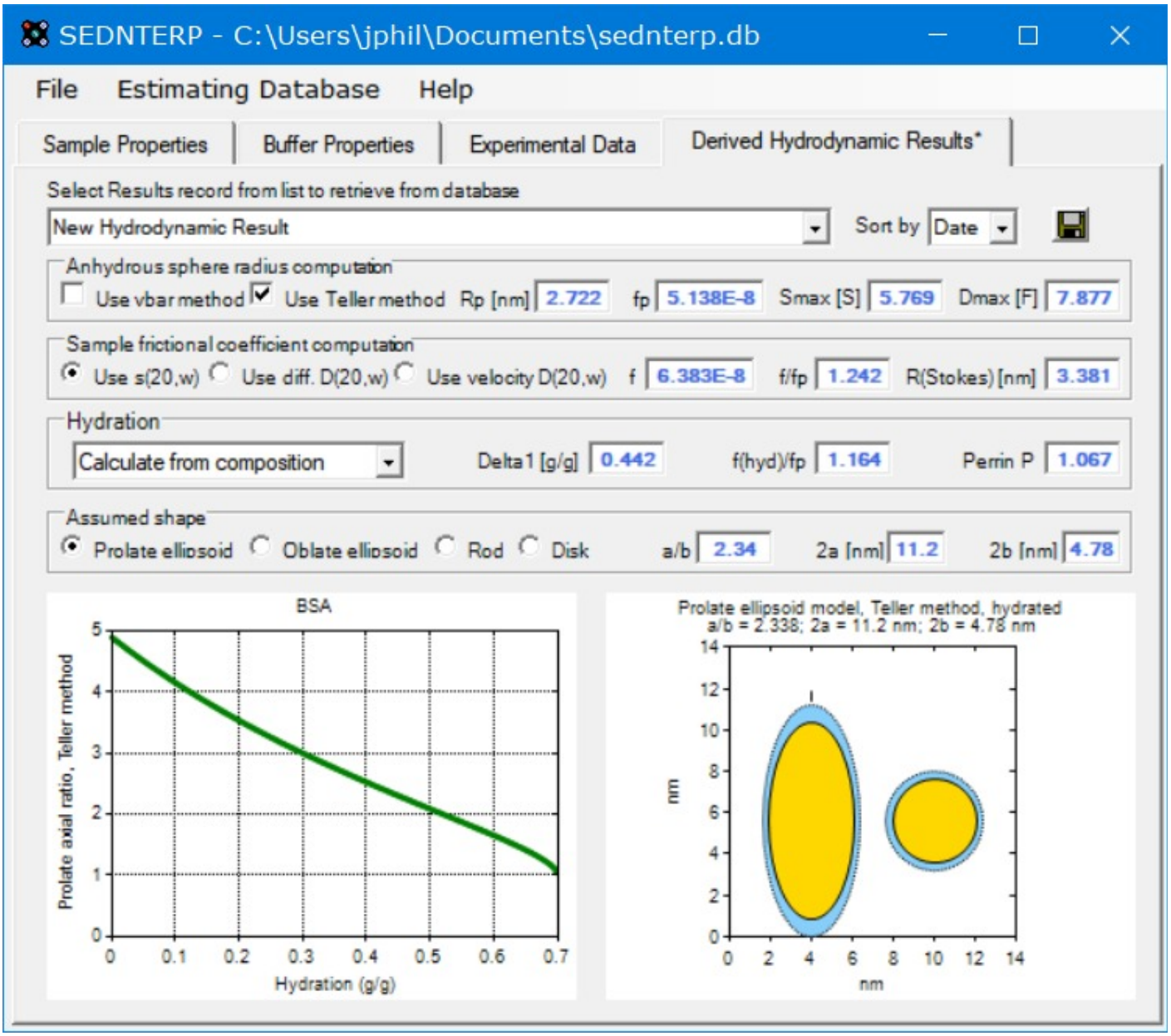
A previously saved Diffusion Experiment record can be retrieved from the database when chosen by name from a drop-down list box located at the upper left in this section. Whether that list is sorted by name, or by the recorded experiment date, is controlled by the **Sort by** drop-down list to its right. The **Save** button at the upper right (with a disk icon) will initiate saving a new Diffusion Experiment record (or over-writing the current one) by bringing up the Saving Data in Databases form.

4.12 Derived Hydrodynamic Results major tab

The Derived Hydrodynamic Results major tab of the Main form calculates and displays multiple parameters from SEDNTERP's hydrodynamic calculations, providing information about molecular hydration and asymmetry. These results are only available (and this tab can only be selected) if enough information to compute a molar mass has been entered, and several of the calculations also require a sedimentation or diffusion coefficient.

To understand what these results mean, and how they are calculated, you should read the topics about Hydration and Asymmetry, Calculating an Equivalent Radius, Calculating the Minimum Frictional Coefficient, Calculating Smax and Dmax, Calculating Maximum Hydration, Calculating Hydration from Asymmetry Estimates, Calculating Ellipsoidal Asymmetry Calculating Cylindrical Asymmetry, Calculating the Stokes Radius, and Limitations on Estimating Hydration and Asymmetry (and the references listed therein for more in-depth descriptions).

The image below shows the Derived Hydrodynamic Results tab loaded with values computed from the '(9/12/1992) BSA sed velocity' experiment record that is stored in the startup database installed with the program.



The major options that affect the hydrodynamic results are the grouped in four panels in the upper half of this tab page, with the entries in each panel affecting the calculations done in the panels below it.

The anhydrous sphere radius computation panel

This top-most panel concerns the equivalent radius, and three quantities that can be derived from that value: the minimum frictional coefficient, the maximum possible sedimentation coefficient, S_{max}, and the maximum possible diffusion coefficient, D_{max}. These are the only hydrodynamic results that can be calculated without an experimental sedimentation

coefficient or diffusion coefficient.

The two check boxes on the left control whether the equivalent radius is calculated by the vbar method or the Teller method. Which of those methods is used by default can be set using the 'Alter program defaults...' item on the File Menu.

It is also possible to manually enter a known equivalent radius value directly into that text box. This 'user override' state is then indicated by neither of the vbar or Teller method check boxes being checked, and also by displaying the radius value as green text.

The sample frictional coefficient computation panel

The three radio buttons on the left side of this panel determine which experimental data should be used for calculating the experimental sample frictional coefficient, *f*. Depending on what experimental data is presently loaded, *f* may be calculated from a sedimentation coefficient, from a diffusion-only experiment, or from a diffusion coefficient derived from a sedimentation velocity experiment.

Once this choice is made, the f/f_0 ratio (ratio of the actual frictional coefficient to the minimum possible value) and the Stokes radius are calculated and displayed.

The hydration panel

The third panel concerns hydration (δ_1), and two other important hydrodynamic parameters that can be calculated once a hydration option is specified and the f/f_0 ratio is known. The first (in display order) of those parameters is the f_{hyd}/f_0 ratio, which compares the frictional coefficient of a hydrated vs. un-hydrated (anhydrous) equivalent sphere (a ratio which therefore depends only on hydration, not molecular shape).

The second of these parameters is the Perrin translational friction function, *P*, which is the ratio of the frictional coefficient for the anhydrous molecule to that for an anhydrous sphere with the same volume. The advantage of the ratio *P* is that it depends *only* on shape or flexibility. Thus together the f_{hyd}/f_0 ratio and the Perrin *P* separate the f/f_0 ratio into a hydration-only portion and a shape-only portion.

The pull-down list on the left side gives 8 hydration options:

1. **Unknown** -- Without any hydration information, only a maximal value for hydration, $\delta_{1,max}$, and maximal values for f_{hyd}/f_0, *P*, and asymmetry can be calculated.
2. **Default** -- This option sets δ_1 to its default value (which is 0.3 g/g when the program is installed, but can be revised by the user).
3. **Use known value** -- This option allows the user to enter a known δ_1 value directly into the

text box to the right.

4. **Calculate from composition** -- This option will use the δ_1 value estimated from the sample's composition.
5. **Calculate from prolate a/b ratio** -- This option will calculate δ_1 based on an assumed prolate ellipsoid shape model and an *a/b* axial ratio that the user enters into the text box within the 'Assumed shape' panel below.
6. **Calculate from oblate a/b ratio** -- This option will calculate δ_1 based on an assumed oblate ellipsoid shape model and an *a/b* axial ratio that the user enters into the text box within the 'Assumed shape' panel below.
7. **Calculate from rod L/d ratio** -- This option will calculate δ_1 based on an assumed rod shape model and a *L/d* length/diameter ratio that the user enters into the text box within the 'Assumed shape' panel below.
8. **Calculate from disk d/H ratio** -- This option will calculate δ_1 based on an assumed disk shape model and a *d/H* diameter/height ratio that the user enters into the text box within the 'Assumed shape' panel below.

The assumed shape panel

This panel contains 4 radio buttons that select one of four shape models:

1. Prolate ellipsoid
2. Oblate ellipsoid
3. Rod (a cylinder with length >= diameter)
4. Disk (a cylinder with diameter >= height)

The asymmetry versus hydration and shape/dimensions graphs

A graph of asymmetry (*a/b*, *L/d*, *or d/H* ratio) as a function of hydration is shown at the bottom left. At the bottom right a picture of the ellipsoid or cylinder model for your molecule is shown on a true length scale. This diagram also shows the hydration layer (note that by tradition it is assumed that the water of hydration expands the molecule uniformly, rather than coating the molecule with a constant thickness).

TIP: Each of these small graphs can be temporarily expanded to the full size of the tab page by clicking with the mouse anywhere within the graph area. Clicking again returns to the smaller size.

Saving or retrieving a Hydrodynamic Result record to/from the database

The **Save to Database** button opens the Saving Data in Databases form. Saving Result records stores the various options on the Results form, along with the experiment, sample, and buffer information, so those can all be retrieved in one operation.

Why should I save a Hydrodynamic Result record? Yes, changing the options on this form yields the new results almost instantaneously, so why bother? The primary reason to save the record is so you can document and reproduce computations that you intend to use in publications or reports.

A previously saved Hydrodynamic Result record can be retrieved from the database when chosen by name from a drop-down list box located at the upper left in this section. Whether that list is sorted by name, or by the date the record was last modified, is controlled by the **Sort by** drop-down list to its right.

4.13 Evaluate ks and Extrapolate s0(20,w) form

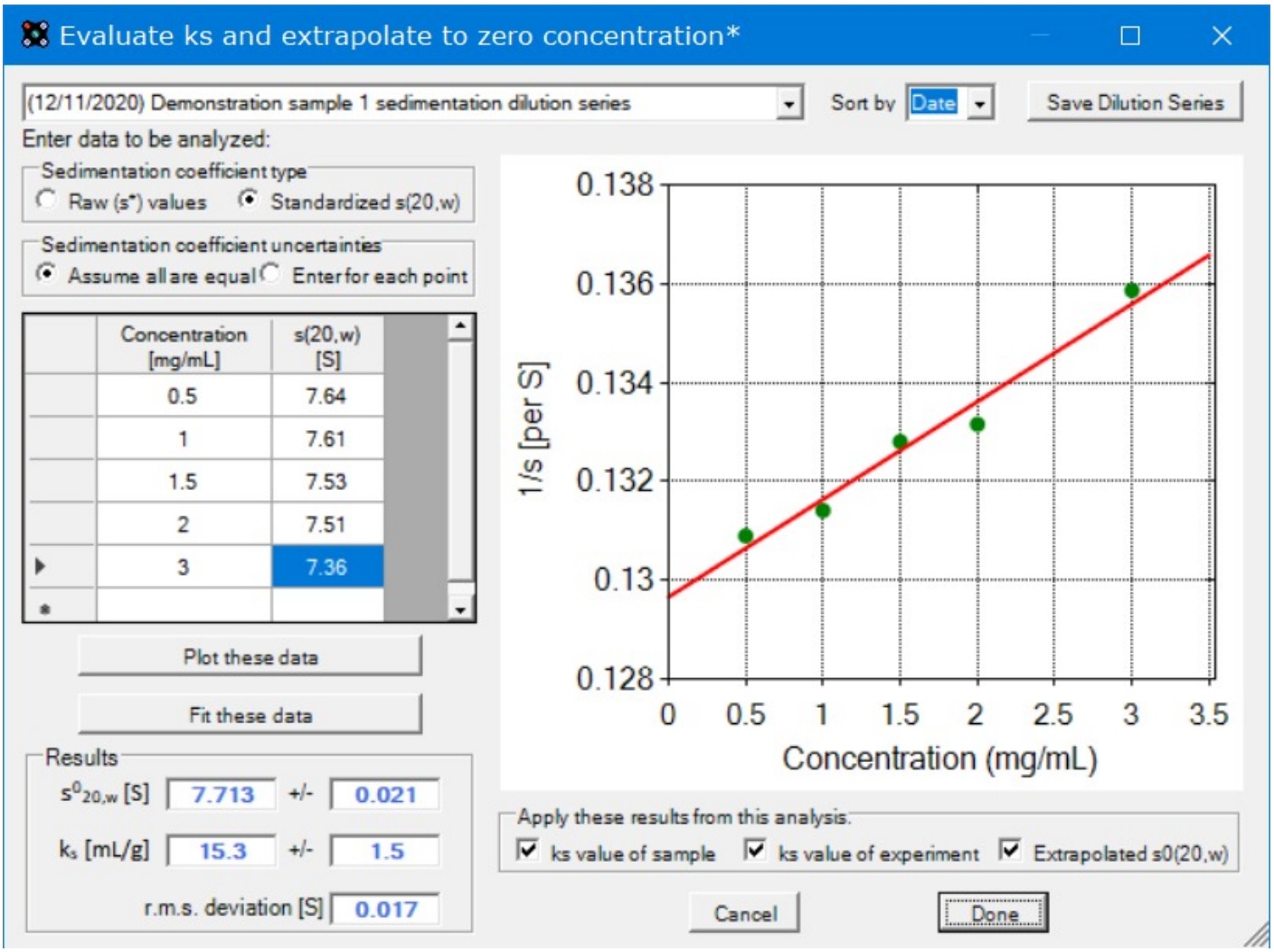
This form allows the user to enter sedimentation coefficients measured at multiple concentrations for the same molecule, and then fit those data to a line (Eq. 34 below) to determine k_s and (optionally) use the extrapolated value at zero concentration as the final $s^0_{20,w}$ value.

Equation 34:

$$\frac{1}{s(c)} = \frac{1}{s_0}\left(1 + k_s \cdot c\right)$$

where $s(c)$ is the sedimentation coefficient (either raw or standardized) at finite weight concentration c, and s_0 the value that would be observed at infinite dilution.

The image below shows this form when loaded with velocity data for 'Demonstration sample 1' (a made-up example that is stored in the startup database installed with the program):



Options about data type and entry

At the upper left are two panels with radio buttons. The top panel is for specifying whether the sedimentation coefficients to be entered in the table below are 'raw' (s*) values or standardized $s_{20,w}$ values. If raw values are used, they will be corrected to $s_{20,w}$ values when the fitting results are reported. When this form is invoked from the Sedimentation Velocity minor tab the conversion to $s_{20,w}$ values uses the $\bar{\nu}$ and buffer density and viscosity taken from that tab page; when it is invoked from the Sample Properties major tab, those properties are irrelevant since no $s^0_{20,w}$ value can be returned.

How to enter, delete, and plot values in the table

The table initially contains only one empty row, but it will automatically add new rows as data are

entered. The value entered in each column must be >= 0. Consequently data entry is easiest if you use the TAB key to sequentially move to the next column or row (hitting an ENTER in the 2nd or 3rd column moves directly down to the next row, leaving an invalid entry to its left).

If you need to delete a row, you cannot do so while it has an empty column, so just enter any positive number. Then select the entire row to be deleted by clicking on its gray (empty) row label on the left (you may have to move the cursor to another row first). Once the entire row is selected it can be deleted using the DELETE key.

Once there are at least 2 complete rows in the table you can plot the current data points in the graph to the right using the **Plot these data** button. Note that if you entered 'raw' sedimentation coefficients those values are plotted as-entered (not as $s_{20,w}$ values).

Fitting the data

Once three valid rows of data are present in the table the **Fit these data** button becomes active. (Obviously two data points are sufficient to define a line, but three are required in order to get an estimate for the uncertainty of the slope and intercept.)

When a fit is completed the best-fit line is shown in the graph as well as the individual data points. After converting to $s_{20,w}$ values if needed, the resultant zero-concentration intercept (the $s^0_{20,w}$ value) and slope (the k_s value) and their estimated standard errors are reported in the panel at the bottom left. Note that the reported r.m.s. deviation of the data points from the best-fit line is calculated after accounting for the inversion in Equation 34.

Returning values for use in other calculations

Check boxes in the 'Apply these results from this analysis' panel below the graph govern whether the new k_s value will be assigned to the current sample and/or the current sedimentation velocity experiment, and whether the new intercept should be used as the $s^0_{20,w}$ value. (The latter two options are only active when this form is invoked from the Sedimentation Velocity minor tab).

Retrieving or saving a Sedimentation Dilution Series record from/to the database

A previously saved Sedimentation Dilution Series record can be retrieved from the database when chosen by name from a drop-down list box located at the upper left in this section. Whether that list is sorted by name, or by the date the Sedimentation Dilution Series record was last modified, is controlled by the **Sort by** drop-down list to its right. The **Save** button at the upper right (with a disk icon) will initiate saving a new Sedimentation Dilution Series record (or over-writing the current one) by bringing up the Saving Data in Databases form.

4.14 Evaluate kD and Extrapolate D0(20,w) form

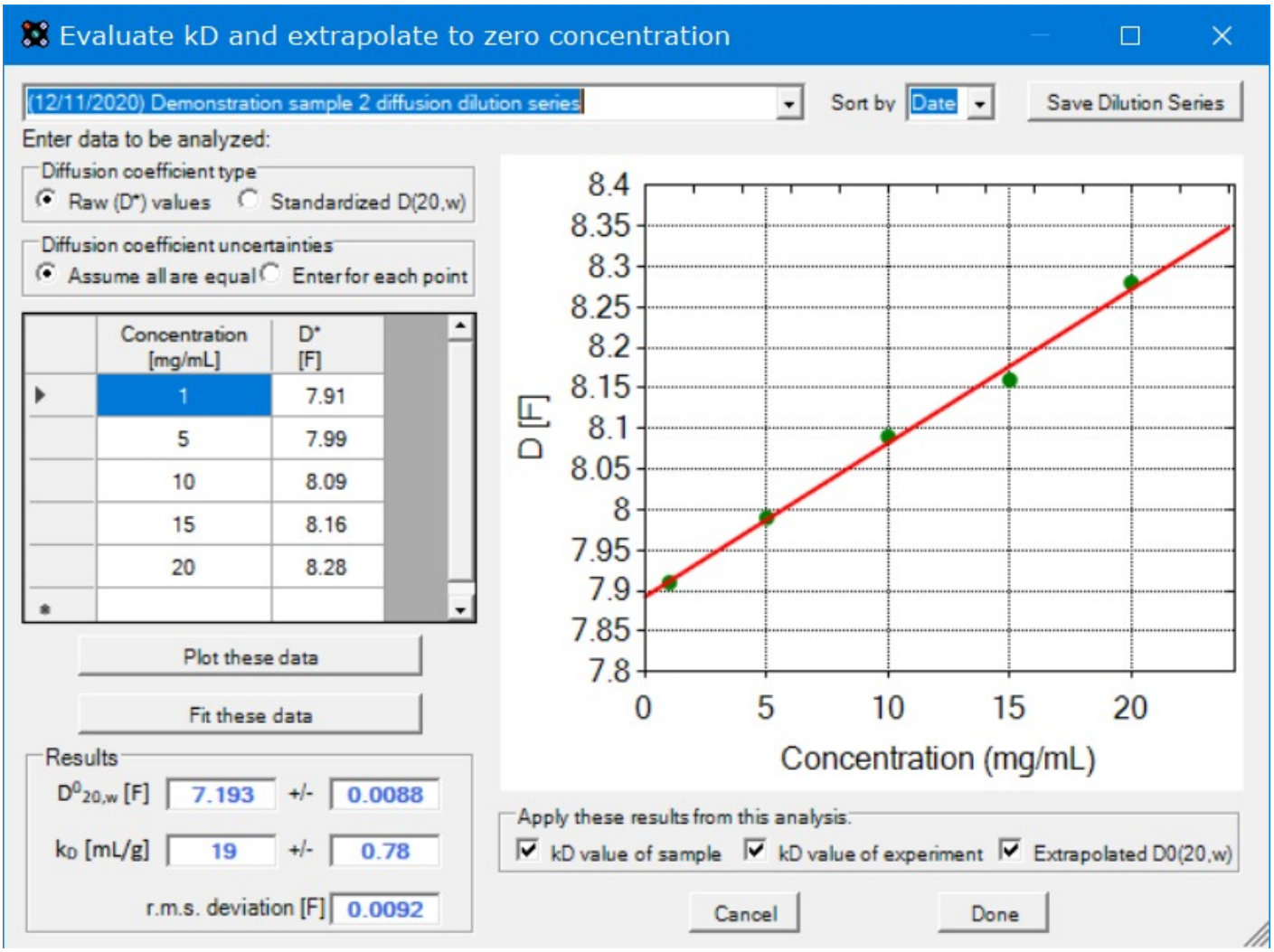
This form allows the user to enter diffusion coefficients measured at multiple concentrations for the same molecule, and then fit those data to a line (Eq. 44 below) to determine k_D and (optionally) use the extrapolated value at zero concentration as the final $D^0_{20,w}$ value.

Equation 44:

$$D(c) = D_0\left(1 + k_D \cdot c\right)$$

where $D(c)$ is the diffusion coefficient (either raw or standardized) at finite weight concentration c, and D_0 the value that would be observed at infinite dilution.

The image below shows this form when loaded with velocity data for 'Demonstration sample 2' (a made-up example that is stored in the startup database installed with the program):



Options about data type and entry

At the upper left are two panels with radio buttons. The top panel is for specifying whether the diffusion coefficients to be entered in the table below are 'raw' (D*) values or standardized $D_{20,w}$ values. If raw values are used, they will be corrected to $D_{20,w}$ values when the fitting results are reported. When this form is invoked from the Diffusion minor tab the conversion to $D_{20,w}$ values uses the buffer viscosity and temperature taken from that tab page; when it is invoked from the Sedimentation Velocity minor tab it takes those properties from there, and when invoked from the Sample Properties major tab the buffer viscosity is taken from the Buffer Properties major tab and the temperature from the Diffusion minor tab.

How to enter, delete, and plot values in the table

The table initially contains only one empty row, but it will automatically add new rows as data are

entered. The value entered in each column must be >= 0. Consequently data entry is easiest if you use the TAB key to sequentially move to the next column or row (hitting an ENTER in the 2nd or 3rd column moves directly down to the next row, leaving an invalid entry to its left).

If you need to delete a row, you cannot do so while it has an empty column, so just enter any positive number. Then select the entire row to be deleted by clicking on its gray (empty) row label on the left (you may have to move the cursor to another row first). Once the entire row is selected it can be deleted using the DELETE key.

Once there are at least 2 complete rows in the table you can plot the current data points in the graph to the right using the **Plot these data** button. Note that if you entered 'raw' diffusion coefficients those values are plotted as-entered (not as $D_{20,w}$ values).

Fitting the data

Once three valid rows of data are present in the table the **Fit these data** button becomes active. (Obviously two data points are sufficient to define a line, but three are required in order to get an estimate for the uncertainty of the slope and intercept.)

When a fit is completed the best-fit line is shown in the graph as well as the individual data points. After converting to $D_{20,w}$ values if needed, the resultant zero-concentration intercept (the $D^0_{20,w}$ value) and slope (the k_D value) and their estimated standard error are reported in the panel at the bottom left. The r.m.s. deviation of the data points from the best-fit line is also reported.

Returning values for use in other calculations

The left-most check box in the 'Apply these results from this analysis' panel below the graph governs whether the new k_D value will be assigned to the current sample. The central check box determines whether the the new k_D value will be assigned to the current diffusion or sedimentation velocity experiment (the experiment from which the form was invoked). The third check box controls whether the new intercept should be used as the $D^0_{20,w}$ value. (The second and third check boxes are only active when this form is invoked from the Diffusion minor tab or the Sedimentation Velocity minor tab).

Retrieving or saving a Diffusion Dilution Series record from/to the database

A previously saved Diffusion Dilution Series record can be retrieved from the database when chosen by name from a drop-down list box located at the upper left in this section. Whether that list is sorted by name, or by the date the Diffusion Dilution Series record was last modified, is controlled by the **Sort by** drop-down list to its right. The **Save** button at the upper right (with a disk icon) will initiate saving a new Diffusion Dilution Series record (or over-writing the current one)

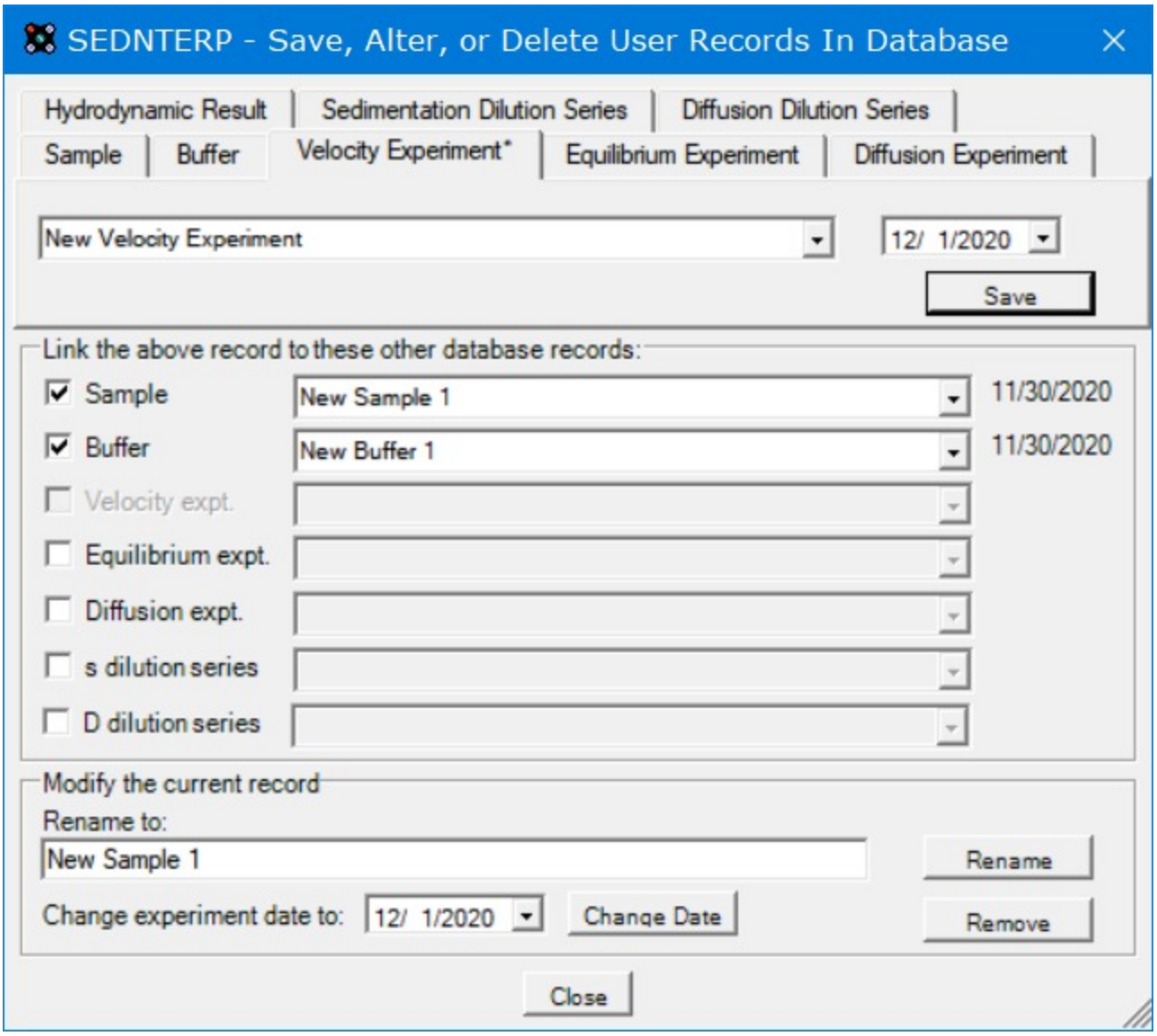
by bringing up the Saving Data in Databases form.

4.15 Saving Data in Databases form

Using this form, Sample, Buffer, Sedimentation Velocity, Sedimentation Equilibrium, Diffusion, Hydrodynamic Result, Sedimentation Dilution Series, or Diffusion Dilution Series records can be stored in the database file. This form also serves as a utility to modify or remove existing records.

Note: SEDNTERP keeps track of whether the properties and parameters associated with the different types of database records have been modified since the record was last saved, and marks which records are 'dirty' and likely need to be saved by adding an asterisk to the name of the corresponding tab page. Those asterisks will also appear on the corresponding tabs of the Saving Data in Databases form (until those records are saved).

The image below illustrates what the form looks like when saving a new Sedimentation Velocity Experiment record. The 8 tabs across the top of the form allow the user to save or modify any of the 'user' records in the database.



The user should supply an appropriate name for the new record to replace the default 'New Velocity Experiment' name in the box at top left. Previously entered record names can be seen (and loaded if you wish to modify their name or date) by pulling down this drop-down list and selecting the desired name.

TIP: Saving a record in the database is similar to saving a word-processing document, but with the record name substituting for the file name. Thus if you want to make a modified copy of a current record, first load it, then make any desired modifications, and just save it under a new name.

TIP: It is obviously important for each user to create an intelligent naming scheme if many experiments, samples, and/or buffers are to be stored in the database. Also, it is may be useful for different users from the same lab to have their experiments in a totally different database file in order to avoid overwriting data records because of repetitive names or human error.

Note that all 3 types of Experiment record require that you enter an Experiment Date using the calendar control at top right (for the other record types the database will automatically maintain a 'last modified' date). The use of an Experiment Date allows for multiple experiments with the same name but different dates. For Experiment records the name + date combination must be unique; for the other record types (Sample, Buffer, *etc.*) the name must be unique.

Creating links to other record types

In the middle of the form there is a panel for specifying any logical links between this new Velocity Experiment record and other types of records. Such "linked" records will then automatically be loaded whenever this new Velocity Experiment record is retrieved from the database. Normally any of the 3 types of Experiment records should be linked to a Sample record (specifying what sample was measured) and also to a Buffer record (containing the buffer properties needed to properly interpret the experiment). In the above example, a Sample record named 'New Sample 1' and a Buffer record named 'New Buffer 1' were currently loaded in the program when saving a new Velocity Experiment was initiated, so the program has automatically set up links to these Sample and Buffer records.

If the molar mass of the sample being measured in this velocity experiment comes from sedimentation equilibrium, then you should probably also link this Velocity Experiment record to that Equilibrium Experiment record.

Similarly, if you are computing this sample's molar mass by combining a sedimentation coefficient from this velocity experiment with a diffusion coefficient from a separate diffusion experiment, you would want to link this new Velocity Experiment record to that Diffusion Experiment record. It also may be appropriate to link to a Diffusion Experiment record if you want to be able to switch back and forth between computing the hydrodynamic results based on sedimentation *vs.* diffusion information (to check for consistency).

It also may be appropriate to link a Velocity Experiment record to a Sedimentation Dilution Series record (different concentrations of this same sample), and/or to a Diffusion Dilution Series record (if the velocity experiment also obtained diffusion information, and data are available for multiple concentrations).

Which types of potential links are available depends on which type of record will be saved (which tab of this form is currently displayed), as summarized below:

- Buffer records, and both types of Dilution Series records, can be linked to, but do not have links.
- Each of the 3 types of Experiment records can link to both of the other types.
- Sedimentation Velocity Experiment records can link to either type of Dilution Series record.
- Diffusion Experiment records can link to a Diffusion Dilution Series record.
- Sedimentation Equilibrium Experiment records cannot link to Dilution Series records.
- Hydrodynamic Result records can link to Sample records and all 3 types of Experiment records, but not to Buffer or Dilution Series records. (Note that a Result record always concerns values at standard conditions, and thus the 'buffer' for a Result record is always water).
- Sample records can link to either type of Dilution Series record.

Saving the new (or modified) record

Once the name, date, and proper links for the new record are entered, it can be saved by simply clicking the **Save** button at the upper right on the form.

Note that this form can also be used to modify (overwrite) current records if you wish to update them. In such situations, after clicking the Save button you will be asked to confirm that you wish to overwrite the current record.

Modifying record names or dates, or removing the current record

Near the bottom of the form is a 'Modify the current record' panel containing controls for removing a record, altering its name, or (for Experiment records only) altering its Experiment Date.

If you wish to modify the name of an existing record, select the record from the drop-down list at the top (which is in alphabetic order). The current name of the record will appear in the top text box. Enter the new name in the **Rename to:** text box. Click the **Rename** button and after confirmation the record name will be immediately modified (this does not require using the **Save** button).

For any of the 3 types of Experiment records the stored Experiment Date can be modified by selecting a new date on the 'Change experiment date to:' calendar control and then clicking the **Change date** button. (This change is immediate, and does not require using the **Save** button.)

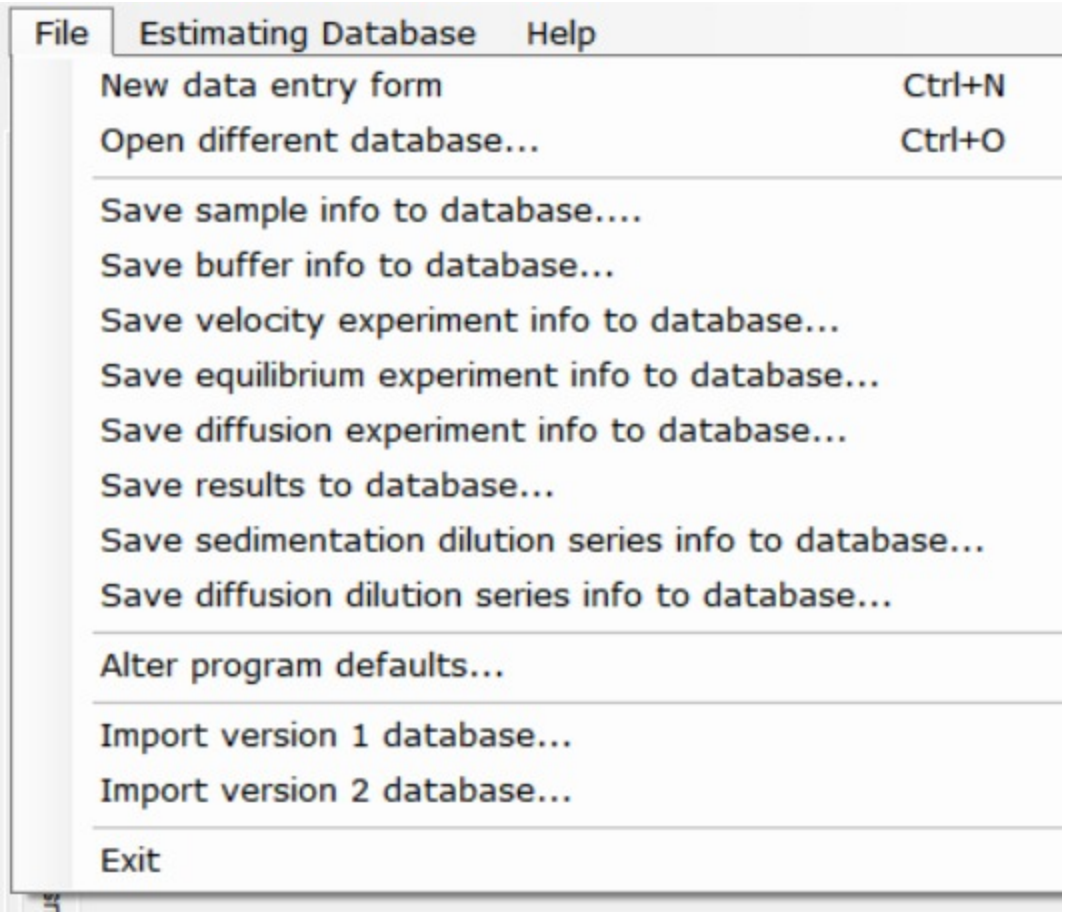
If you wish to remove a record, choose it from the drop-down list box and then press the **Remove** button. After confirmation, the record will be marked as inactive (to preserve database integrity it is not actually deleted, however SEDNTERP provides no mechanism to re-activate such records).

Returning to the main form

When you are done saving or modifying records click the **Close** button at the bottom to return to the Main form.

4.16 File Menu

The **File Menu** on the main form (shown below) contains the following options:



1. **New data entry form** This choice re-initializes the main form. Choose this option to start over with a new sample and buffer, a new set of 3 experiments, and new (empty) derived hydrodynamic results. It will also re-initialize values to those of the program defaults. This command can also be invoked with the shortcut key combination CTRL+N (used for File...New in many Windows applications).

2. **Open different database...** This choice opens a file dialogue window to allow you to choose a new database for saving experimental, sample, and buffer data. The standard name for this database is sednterp3.db. Users can create their own databases by downloading a copy of the startup database from the program website and then re-naming it as desired (while keeping

the .db extension). This command can also be invoked with the shortcut key combination CTRL+O (used for File...Open in many Windows applications).

Note that during installation the database file is placed within a 'Sednterp3' sub-folder of the user's Application Data folder (called AppData in Windows 10). The full path to that folder is typically c:\users\(your user name)\AppData\Roaming\Sednterp3. Note that this AppData folder is automatically backed up by some (but not necessarily all) data backup software, so you may wish to confirm that your database file will be included in your backups.

After opening the new database you will be asked whether you wish to automatically open this database for future program starts.

Saving user records to the database

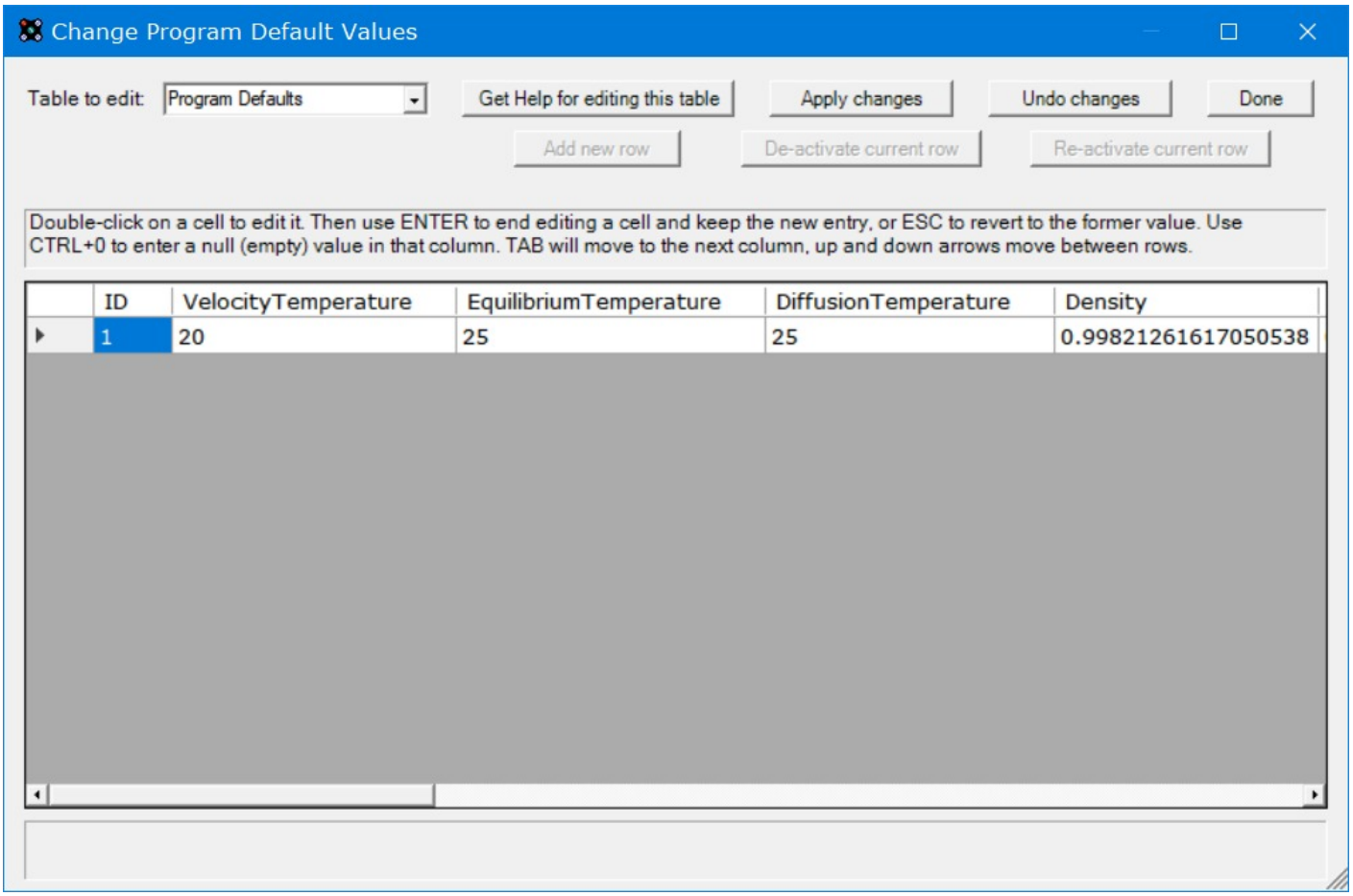
3. **Save sample Info to database...**
4. **Save buffer info to database...**.
5. **Save velocity experiment Info to database...**
6. **Save equilibrium experiment Info to database...**
7. **Save diffusion experiment Info to database...**
8. **Save results info to database...**
9. **Save sedimentation dilution series info to database...**
10. **Save diffusion dilution series info to database...**

All of these choices open the Saving Data in Databases form, but to a different specific page about that particular type of database record (a particular table in the database). The Saving Data in Databases form provides a common interface for saving any part of the data the user has entered or computed. A more detailed description of the different tables in the database is given here.

Changing program default values

11. **Alter program defaults...**

The values for program defaults (also known as User Preferences) are actually stored in a table in the program database (a table that has only 1 row). Therefore this menu choice opens the Edit Physical Properties or Program Defaults form where the user can change the default values. The various default values are detailed here, and the form image is shown below:



Any change in default values takes effect only when the program is reinitialized by being restarted or by the New Data Entry choice of this menu.

Importing an old database from SEDNTERP versions 1 or 2

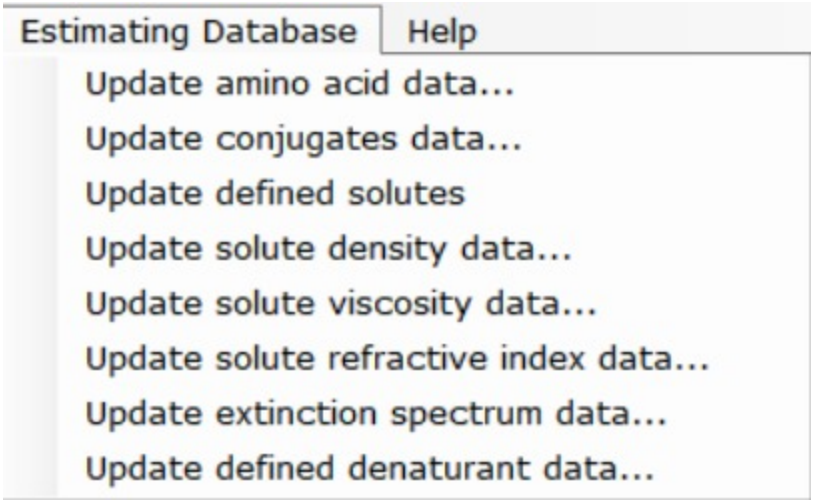
12. **Import version 1 database...**

13. **Import version 2 database...**

The above two menu items will initiate importing the database used by a previous version of SEDNTERP. Versions 1.x used two Microsoft Access format databases (separate files for the user's experimental data and the physical constants data), while version 2 used a single SQLite database file. SEDNTERP version also uses a single SQLite database file, but it uses different table structures and additional tables. More details about importing old databases are found in the Importing Databases from Earlier Versions topic.

14. **Exit** This choice will exit the program. If any information about the current sample, buffer, or experiments has been changed but not saved, you will be asked whether you wish to save those altered records before exiting.

4.17 Estimating Databases Menu

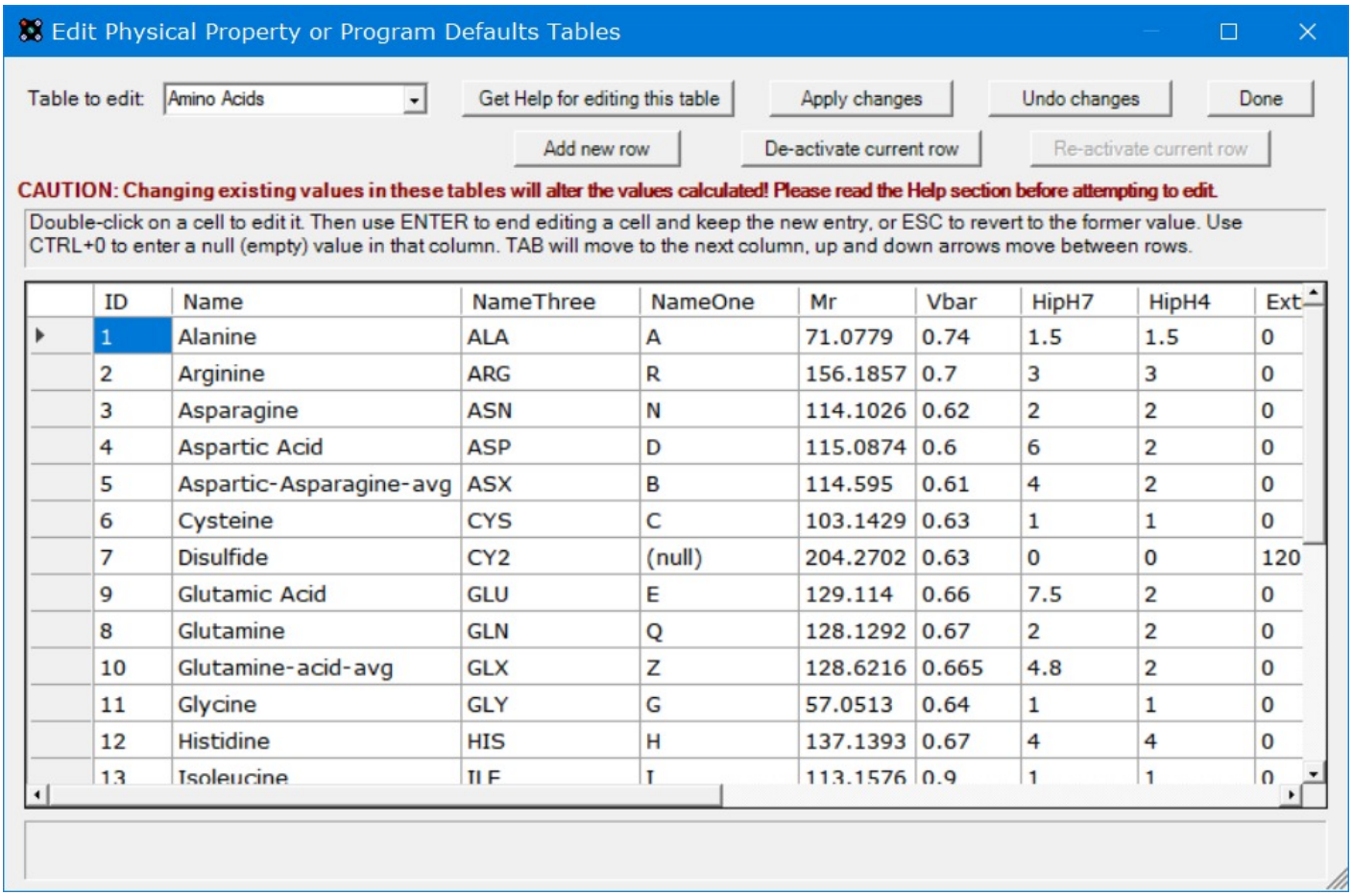


This menu allows the user to access the Physical Constants tables within the program database to modify or add to the calculation/interpolation data used by SEDNTERP.

CAUTION: Altering these physical constant values will change the properties of samples or buffers that SEDNTERP calculates, and/or the derived hydrodynamic properties of the molecule, and therefore should be done with great caution and care, and after studying the details about the table entries contained in separate topics of this document (listed under 'See Also' for this topic).

If you have questions regarding how to add to or modify entries in these Physical Constants tables, please contact the authors. Moreover, if you believe you have found errors in a database that you wish to correct in your own copy, the programmers ask that you please contact us so that we may publish the fixes to all interested parties.

The two things most likely to be modified by users are adding new conjugates and/or unusual amino acids. Those tables, as well as data about buffer components, defined denaturants, and extinction spectra can all be altered using the **Edit Physical Property or Program Defaults Tables** form. This form is shown below as it looks when editing the amino acids table.



ID	Name	NameThree	NameOne	Mr	Vbar	HipH7	HipH4	Ext
1	Alanine	ALA	A	71.0779	0.74	1.5	1.5	0
2	Arginine	ARG	R	156.1857	0.7	3	3	0
3	Asparagine	ASN	N	114.1026	0.62	2	2	0
4	Aspartic Acid	ASP	D	115.0874	0.6	6	2	0
5	Aspartic-Asparagine-avg	ASX	B	114.595	0.61	4	2	0
6	Cysteine	CYS	C	103.1429	0.63	1	1	0
7	Disulfide	CY2	(null)	204.2702	0.63	0	0	120
9	Glutamic Acid	GLU	E	129.114	0.66	7.5	2	0
8	Glutamine	GLN	Q	128.1292	0.67	2	2	0
10	Glutamine-acid-avg	GLX	Z	128.6216	0.665	4.8	2	0
11	Glycine	GLY	G	57.0513	0.64	1	1	0
12	Histidine	HIS	H	137.1393	0.67	4	4	0
13	Isoleucine	ILE	I	113.1576	0.9	1	1	0

The drop-down list box at the upper left controls whether the amino acid, conjugates, solutes, solute density data, solute viscosity data, solute refractive index data, defined denaturant, extinction, or program defaults table is loaded and displayed for editing. The entire table is loaded into the spreadsheet below (all current rows), and as different tables are selected the number and headings for the columns change.

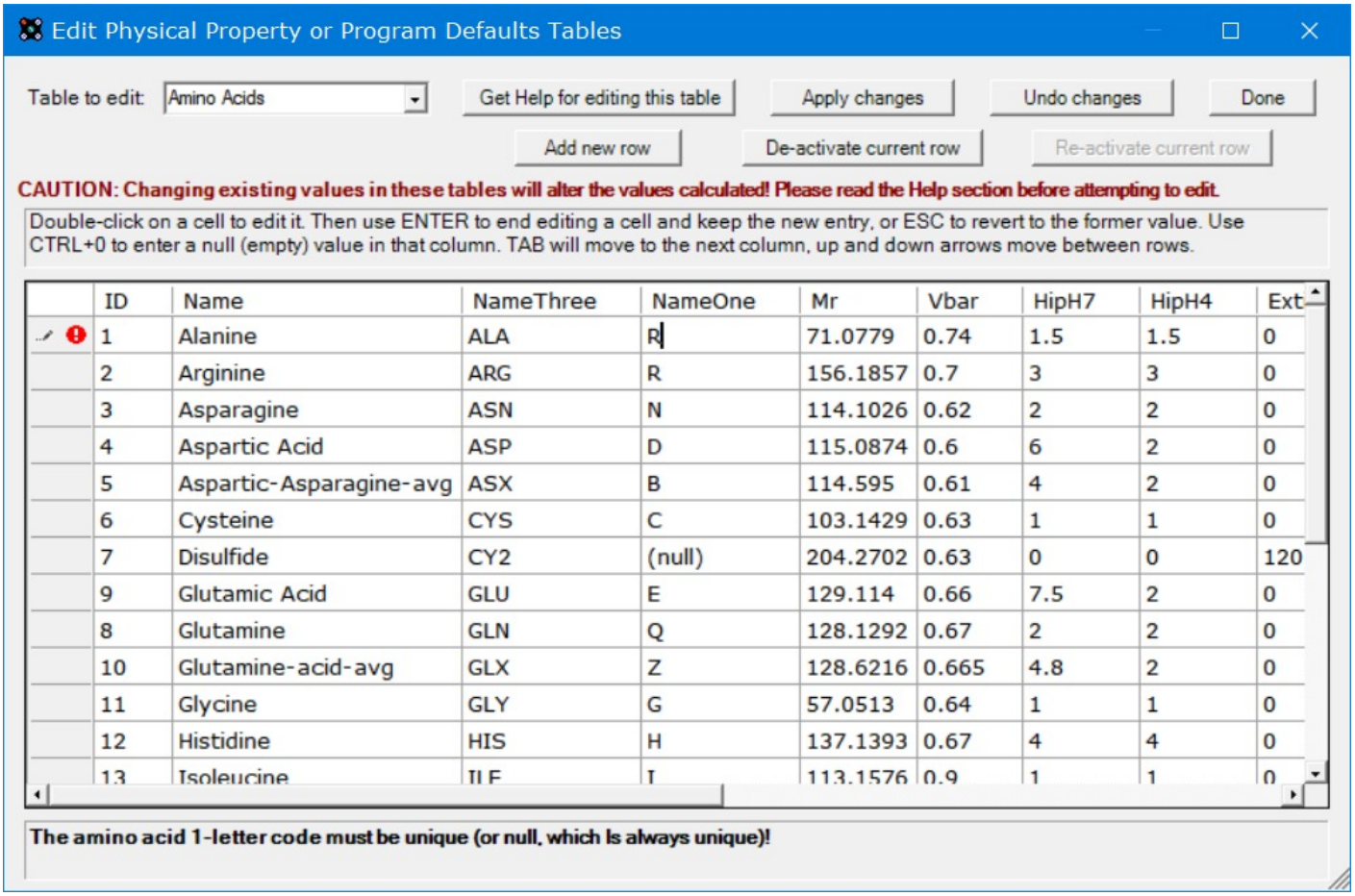
Brief instructions for how to edit individual current entries (cells) are shown in a box above the spreadsheet. Double-click on a cell to begin editing it (while you are editing, the triangle symbol shown in the fixed gray column at the far left changes to a pencil icon). Then use ENTER to finish editing and keep the new entry, or use the ESC key to revert to the former value. To enter a *null* value (empty entry) use the CTRL+0 key combination. Navigate to the next or preceding column using TAB and Shift+Tab, and to the next or preceding row using the down-arrow or up-arrow keys.

No changes are actually written to the database until you click the **Apply changes** button. The **Undo changes** button will un-do all changes made since the Apply changes button was last used (or since this table was loaded if no changes have been written in this session).

Data entry validation

The new or revised entries for some columns must meet certain conditions or they will be flagged as errors. For example, in the amino acids table the 'NameThree' and 'NameOne' column entries (the 3-letter and 1-letter codes for this amino acid) must be unique (they cannot duplicate the codes for another amino acid). Similarly the molar mass (Mr) and vbar entries must exist for every amino acid (these values cannot be *null*).

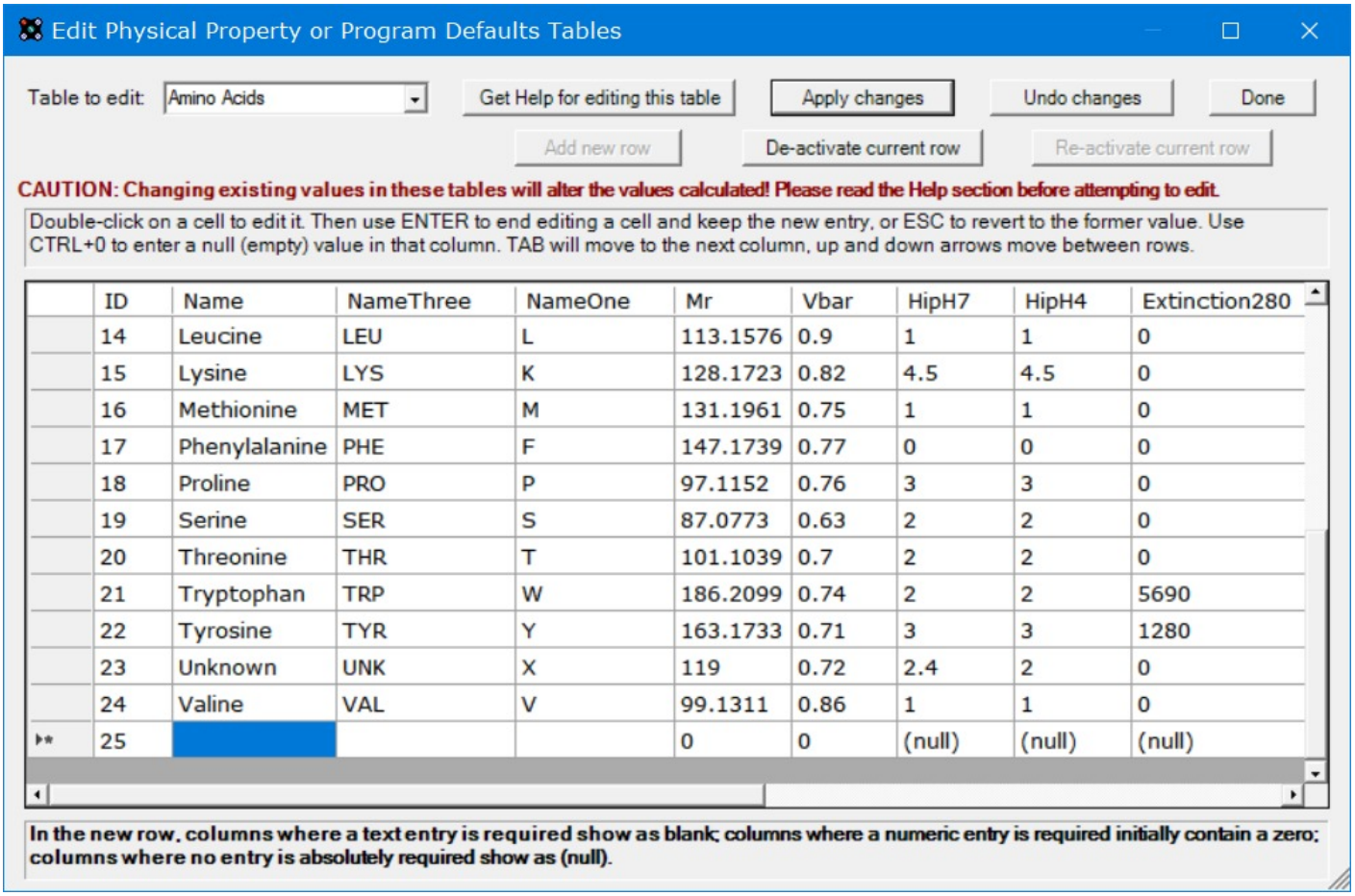
The image below shows an example of what happens when an entry is flagged as an error. In this case the 1-letter code for Alanine was entered as R, which clearly is not unique. As soon as the user attempts to complete that edit, a red exclamation point icon appears in the fixed gray column at the far left (in the row with the error), and the program will not let the cursor leave the cell with the error. An explanatory error message is shown in the panel below the spreadsheet: "The 1-letter amino acid code must be unique (or null, which is always unique)!". That error message can also be seen if you hover the mouse over the red exclamation point.



Adding new rows to the table

New records (new rows) can be created with the **Add new row** button, which will add the new row as the *last* row in the table, as shown below. Here again some instructions are displayed in the

panel below the spreadsheet.



ID	Name	NameThree	NameOne	Mr	Vbar	HipH7	HipH4	Extinction280
14	Leucine	LEU	L	113.1576	0.9	1	1	0
15	Lysine	LYS	K	128.1723	0.82	4.5	4.5	0
16	Methionine	MET	M	131.1961	0.75	1	1	0
17	Phenylalanine	PHE	F	147.1739	0.77	0	0	0
18	Proline	PRO	P	97.1152	0.76	3	3	0
19	Serine	SER	S	87.0773	0.63	2	2	0
20	Threonine	THR	T	101.1039	0.7	2	2	0
21	Tryptophan	TRP	W	186.2099	0.74	2	2	5690
22	Tyrosine	TYR	Y	163.1733	0.71	3	3	1280
23	Unknown	UNK	X	119	0.72	2.4	2	0
24	Valine	VAL	V	99.1311	0.86	1	1	0
25				0	0	(null)	(null)	(null)

Note that the "ID" column (the first column in every table) cannot be edited by the user. The database automatically generates an ID# for newly-added records. Another column that cannot be edited is the "Modified" column (usually the last one in each table), which contains the date and time that record was last modified (in Universal Time).

Sorting the table

When a table is first displayed it is sorted by the Name column, but the user can sort based on any column simply by clicking on the column name at the top. The sort order can be changed from ascending to descending by repeatedly clicking on the column name.

Inactivating or re-activating rows in the table

To preserve database integrity individual records (rows) should never be deleted (*e.g.* with an external program). However it is possible to mark a record as "inactive" by clicking the **De-activate current row** button. Such inactive records will *not* be included in the lists of amino

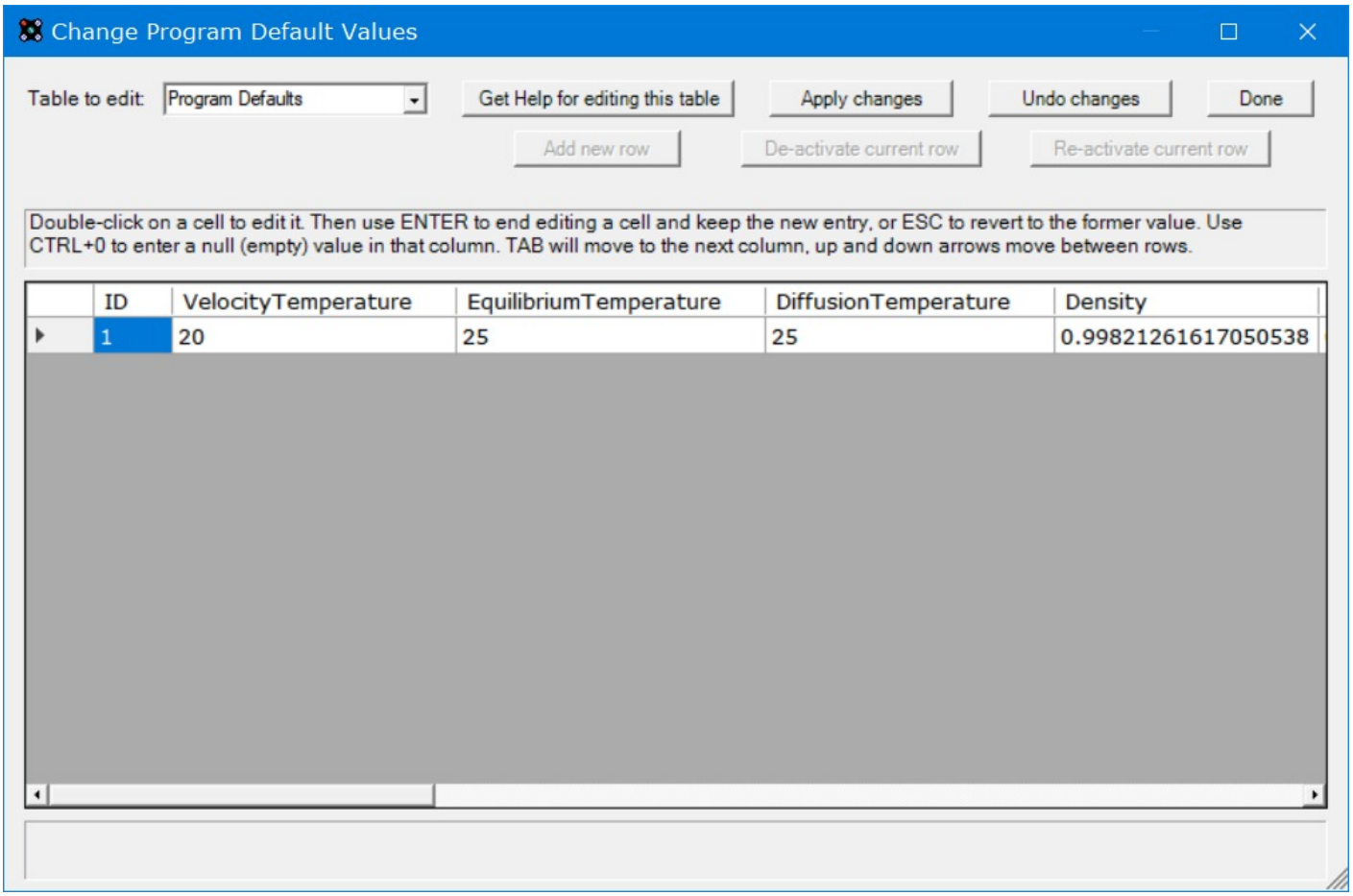
acids or conjugates shown on the Sample Composition form, or included in the buffer solute names or data shown on the Buffer Composition form, but are displayed on this database editing form.

Inactive records can also be re-activated by clicking the **Re-activate current row** button. That button is only active when the current row is itself marked as inactive.

4.18 The Program Defaults Table

SEDNTERP provides default values for nearly every sample or buffer property (with the exception of sample molar mass). Those default values can be changed to the user's preference. For example, if you commonly use a particular buffer, you could set its density, viscosity, and refractive index as the default values.

The default values are stored in a special table in the program database (a table that has only a single row). To modify those values, you can either invoke the **Alter program defaults...** command on the File Menu, or use any of the commands on the Estimating Databases Menu and then change the table being edited to the Defaults table. This will bring up a form as shown below:



The entries in that table can then be edited by double-clicking on the corresponding cell. All of the columns can be seen by using the slider at the bottom of the spreadsheet. (The 'ID' and 'Modified' columns cannot be changed by the user.)

What do the column headings mean?

The following list gives the meaning of the various column headings. The initial default values (on program installation) are in parentheses.

1. **Velocity Temperature** -- the temperature of the sedimentation velocity experiment (20 ^{o}C)
2. **EquilibriumTemperature** -- the temperature of the sedimentation equilibrium experiment (25 ^{o}C)
3. **DiffusionTemperature** -- the temperature of the diffusion experiment (25 ^{o}C)
4. **Density** -- the density of the buffer (0.998213, the density of pure water at 20 ^{o}C)
5. **Viscosity** -- the viscosity of the buffer (0.010016 poise, the viscosity of pure water at 20 ^{o}C)
6. **Vbar** -- the $\bar{\nu}$ of the sample (0.73 mL/g, a good guess for globular proteins)

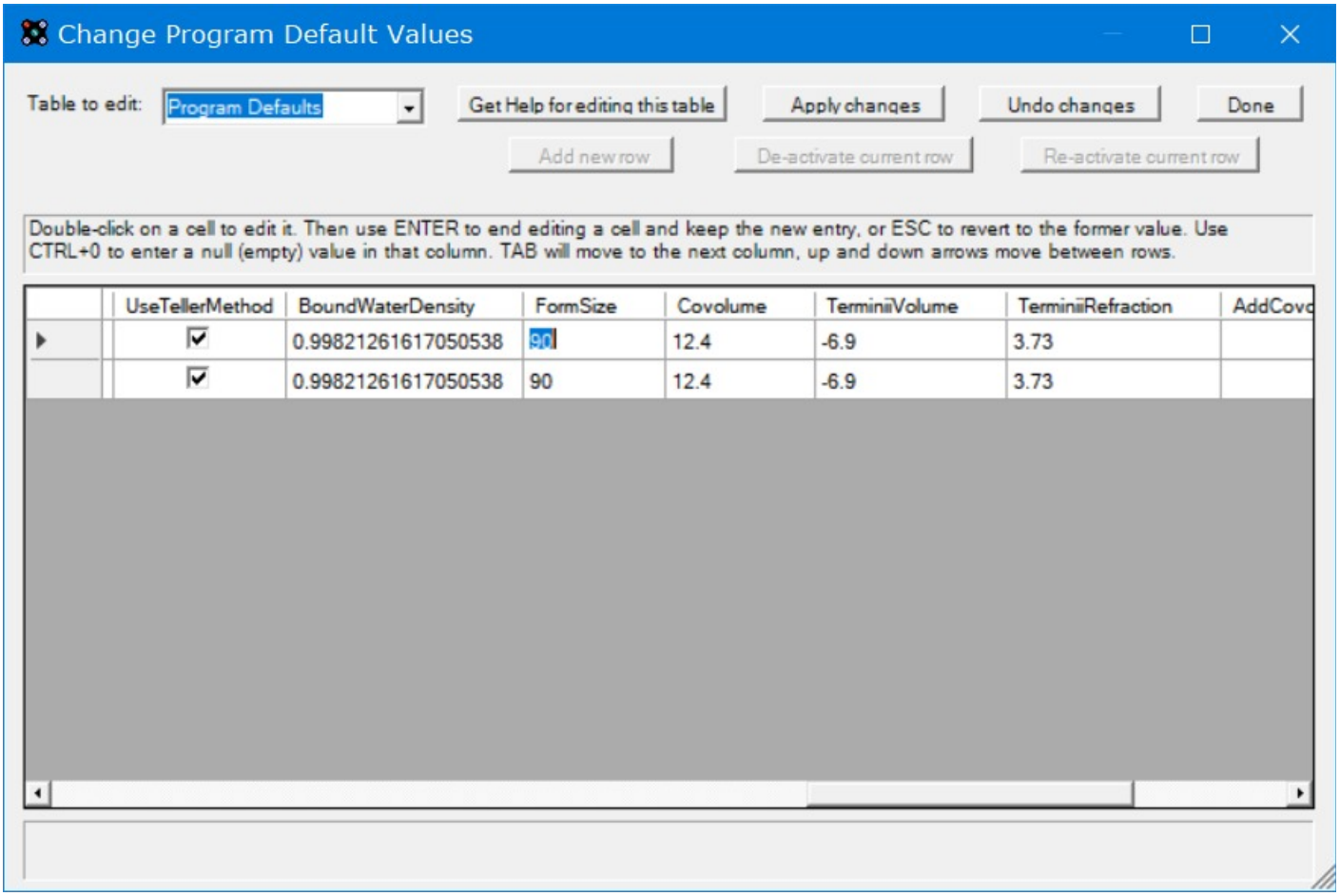
7. **pH** -- the pH of the buffer (7.0)
8. **Ks** -- the concentration dependence of the sedimentation coefficient, k_s (9 mL/g, a good guess for compact globular proteins)
9. **KD** -- the concentration dependence of the diffusion coefficient, k_D (0.9 mL/g, a 'not unreasonable' guess for compact globular proteins at physiological ionic strength, when measured by methods involving a concentration gradient). This default value is definitely the most uncertain, and the rationale for choosing this value can be found here.
10. **RefractiveIndex** -- The refractive index of the buffer (1.333359 at 20 ^{o}C and 589.26 nm, the value for pure water)
11. **RefrIndexWavelength** -- the default wavelength to use when calculating the buffer refractive index (589.26 nm)
12. **DnDcWavelength** -- the default wavelength to use for calculating the *dn/dc* value used in experiments (675 nm)
13. **Hydration** -- the hydration, δ_1 (0.3 g/g, a commonly-used 'typical' value for globular proteins)
14. **pKaAmino** -- the pK_a value for the amino terminus (8.00)
15. **pKaCarboxy** -- the pK_a value for the carboxyl terminus (3.67)
16. **UseTellerMethod** -- a Boolean value that controls the default method for computing the Equivalent Radius. If checked (True) the Teller method is the default; if un-checked (False) the vbar method is the default (Teller)
17. **BoundWaterDensity** -- the default density for the bound water of hydration (0.998213, the density of pure water at 20 ^{o}C)
18. **FormSize** -- the default relative size for the program forms, as a percentage of the nominal size (100)
19. **Covolume** -- the default value for the covolume term that can be optionally used when Calculating Vbar Estimates (12.4 mL/mole)
20. **TerminiiVolume** -- the default value for the summed volume of a polypeptide C- and N-terminus, a value that can be optionally used when Calculating Vbar Estimates (-6.9 mL/mole)
21. **TerminiiRefraction** -- the default value for the summed molar refraction of a polypeptide C- and N-terminus, a value that can be optionally used when Calculating Refractive Increment Estimates (3.73)
22. **AddCovolumeByDefault** -- a Boolean value; if checked (True) the covolume term (item 20 above) will be added by default when calculating $\bar{\nu}$ from composition for new samples (True)
23. **AddTerminiiByDefault**-- a Boolean value; if checked (True) the volume and refraction of the terminal groups (items 21 and 22 above) will be included by default when calculating $\bar{\nu}$ and *dn/dc* from composition for new samples (True)

Saving your changes

Any changes will not actually be written to the database until the **Apply changes** button is clicked. Any change in default values takes effect only when the program is reinitialized by being restarted or by the New Data Entry choice of the File Menu.

The **Undo changes** button will revert to the previous values.

Note that after you click the Apply Changes button the spreadsheet will appear to have added a second, identical row. The image below illustrates what the form looks like after the default size of the program forms is changed from 100% to 90%.



This display of a second row is merely a quirk of this spreadsheet control---there is really still only one row.

Note that the **Add new row**, **De-activate current row**, and **Re-activate current row** buttons are all inactive (those operations do not apply to the Defaults table).

4.19 Importing Databases from Earlier Versions

In this topic matters that are common to importing from either a SEDNTERP version 1.x or version 2 database will be discussed first, followed by matters that are specific to version 1.x or version 2 databases. Before reading further, you should probably read Description of the Databases to learn what information is in the 'physical constants' and 'user' portions of the program database.

Should I import the physical constants portion of the old database?

The short answer is "most probably not", and there are drawbacks to importing certain old values (which will be discussed later). However it is important to think about this issue before you import your old database (and the same considerations apply whether your old database is from SEDNTERP versions 1 or 2).

There are essentially 3 different situations, depending on how you used the program in the past:

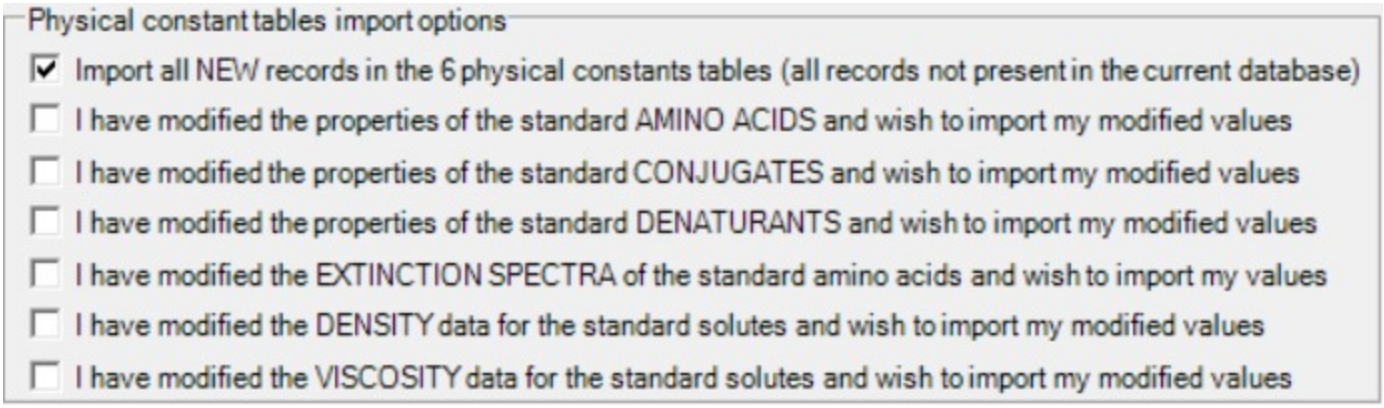
1. You are a typical user who never added new amino acids, new conjugates, new density or viscosity data for buffer components, nor ever altered any of the values for the existing entries in the physical constants tables.
2. You are a more sophisticated user, and needed to add unusual amino acids, additional conjugates, and/or additional density or viscosity data for buffer components. Therefore you want to import those additions to version 3.
3. You are a 'power' user who wanted total control, and were using (for example) a different set of $\bar{v}$ and/or pK_a values for the amino acids. Therefore you want to import those modified values into version 3.

What are the drawbacks to simply importing ALL the old physical constant entries? The primary drawbacks to importing the old amino acid property table are:

1. The SEDNTERP 3 database contains newer, "better" values for the pK_a of the amino acid residues and the amino and carboxy terminus (from Thurkill *et al.* 2006, ref. 93).
2. The SEDNTERP 3 database contains more accurate values for the amino acid residue masses (the values used in current mass spectroscopy software).
3. Similarly some of the conjugate molar masses have been updated in SEDNTERP 3.

How do I control which (if any) physical constant data get imported?

The options for physical constant data import are controlled by the set of check boxes shown below:

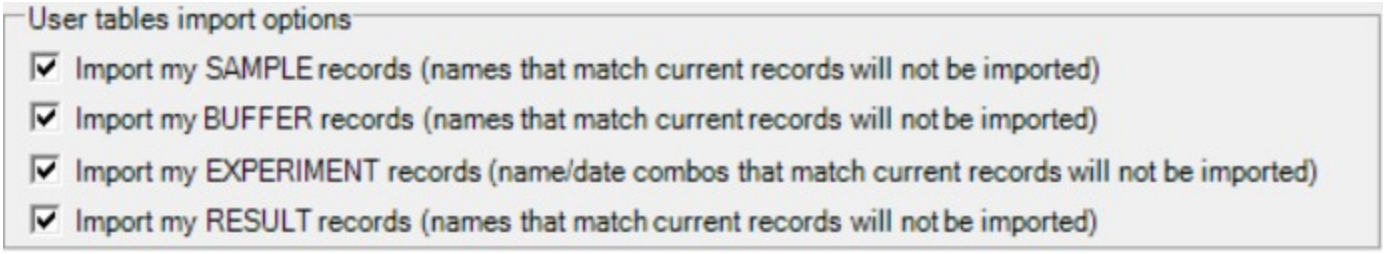


The choices shown above should work for both the 'typical' and 'sophisticated' users (situations 1 and 2 above). The 'typical' user won't actually have any new records that will get imported, but leaving this 'Import all new records' option checked does no harm, and ensures nothing will be lost even if you have forgotten that you added something new.

The other 6 check boxes allow the 'power' user to import only the physical constants tables that have been modified. Again, it is not recommended that you check any of these 6 boxes unless you are sure you want to over-ride the current 'preferred' values.

How do I control which 'user' data get imported?

The options for importing the 'user' or 'experimental' portions of the old database are controlled by the set of check boxes shown below:



By default all of these check boxes are checked, which means all the user records will be imported. Note however that records whose names (or name + date combination for experiment records) match ones already present in the version 3 database will not be imported.

Should I also important my old defaults (user preferences)?

Probably not. The problem with importing your old defaults is that in the ~25 years since SEDNTERP was created there have been changes in the "best" values for the density and viscosity of water. SEDNTERP 3 has been updated to use the current "recommended" values for those properties. The difference in density is only in the sixth decimal place, but for viscosity the difference is 0.04%. Thus if you import your old defaults they will disagree slightly from the values

that SEDNTERP 3 (as installed) will compute.

Be smart! If you have already added new samples, buffers, *etc.* to your SEDNTERP version 3 database, please make a backup copy of that file (usually the file sednterp3.db in the folder c:\users\(your user name)\AppData\Roaming\Sednterp3) before you import your old database, just in case something goes horribly wrong.

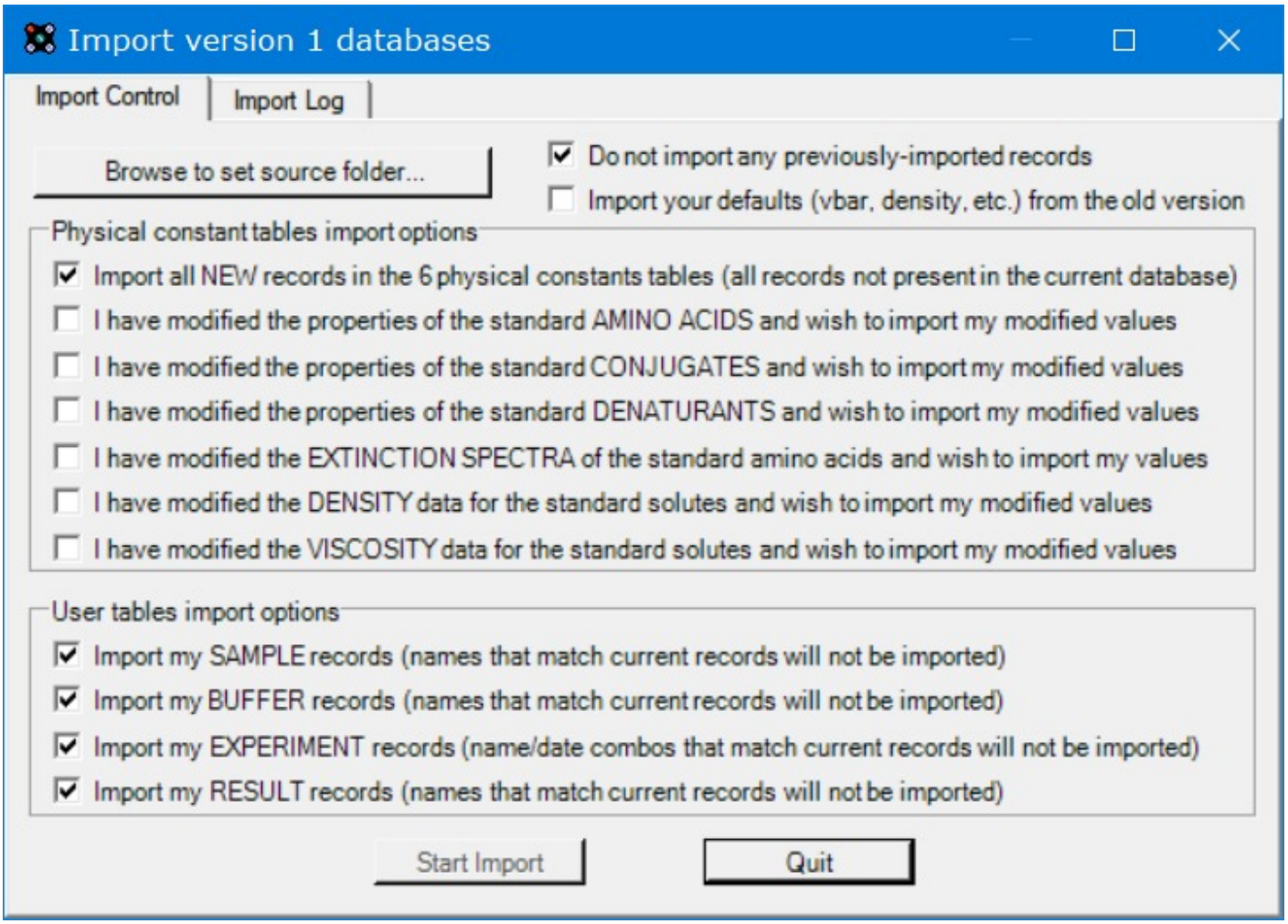
Details about importing from SEDNTERP version 1 versus version 2

The details about importing databases from SEDNTERP versions 1.x and version 2 are given in separate drop-down expandable content sections below. Click on the blue heading or the triangle to the left to expand the section.

Importing a database from SEDNTERP versions 1.x

SEDNTERP versions 1.00 - 1.11 used two Microsoft Access format database files. One database file contains the 'user' or 'experimental' portion (the file sednterp.mdb). A second database file (phyconst.mdb) contains the physical constants data.

Invoking the **Import version 1 database...** command from the File Menu will bring up this form:



The **Browse to set source folder...** button must be used to designate the folder where the version 1 database files are located. (Since the file names are fixed, only a folder needs to be designated). Once the folder is set the **Start import** button becomes active.

The options for controlling the import of the physical constants and user tables, as well as your old defaults, were already discussed above, and should be reviewed before you click the **Start import** button.

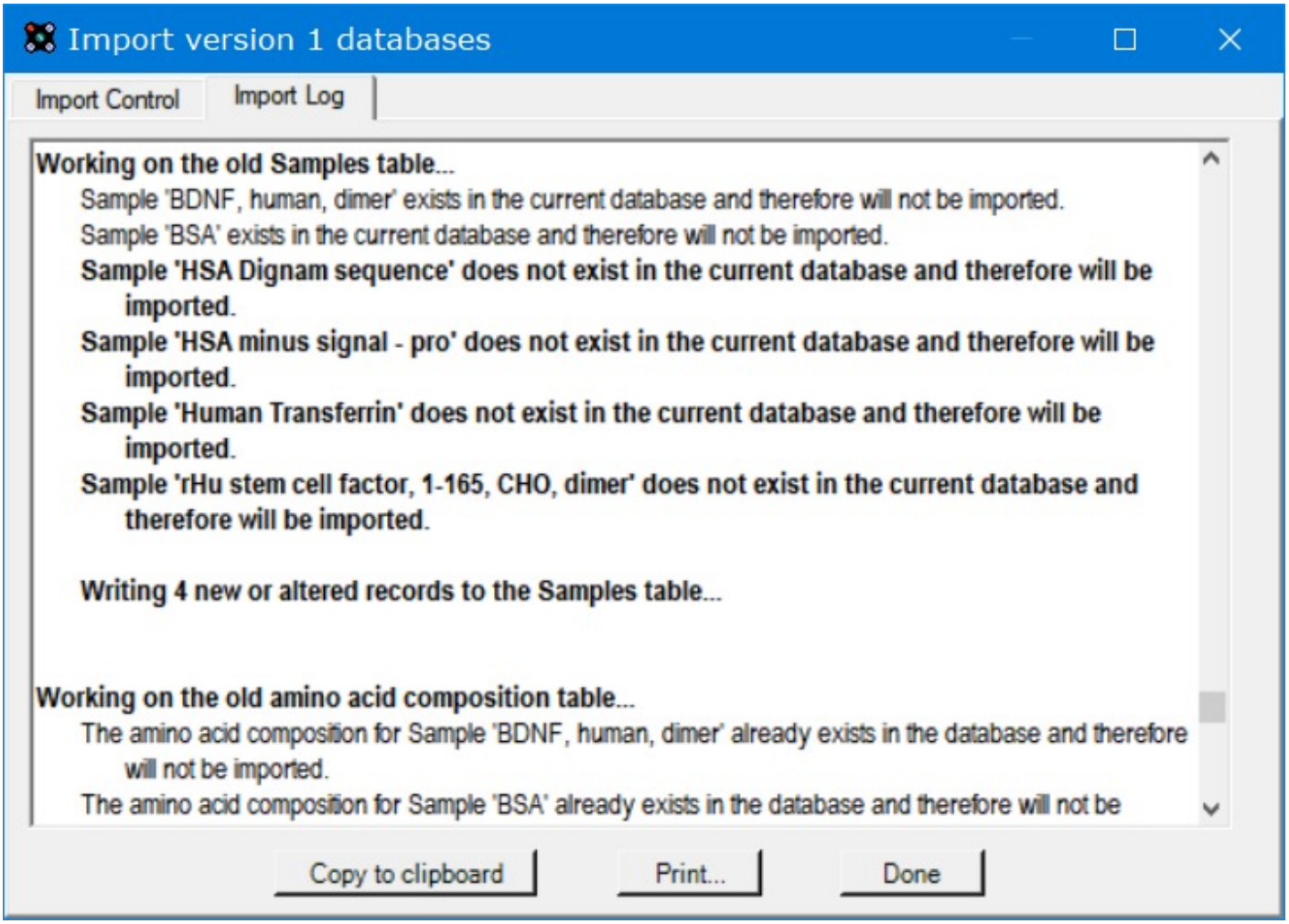
The sole remaining option is the 'Do not import any previously-imported records' check box at top right, which is on by default. As SEDNTERP 3 imports your old databases it marks the records it imports. Therefore if for some reason you wish to go through the import process again, ordinarily it would skip over those previously-imported records, but you can override that behavior by un-checking this box.

NOTE: SEDNTERP 2 used the same mechanism to mark records that it imported from version 1 databases. Therefore if you ever previously imported your version 1

databases into version 2, you should probably un-check 'Do not import any previously-imported records'.

The **Quit** button will simply terminate the import, close this import control form, and return to the Main Data Entry form.

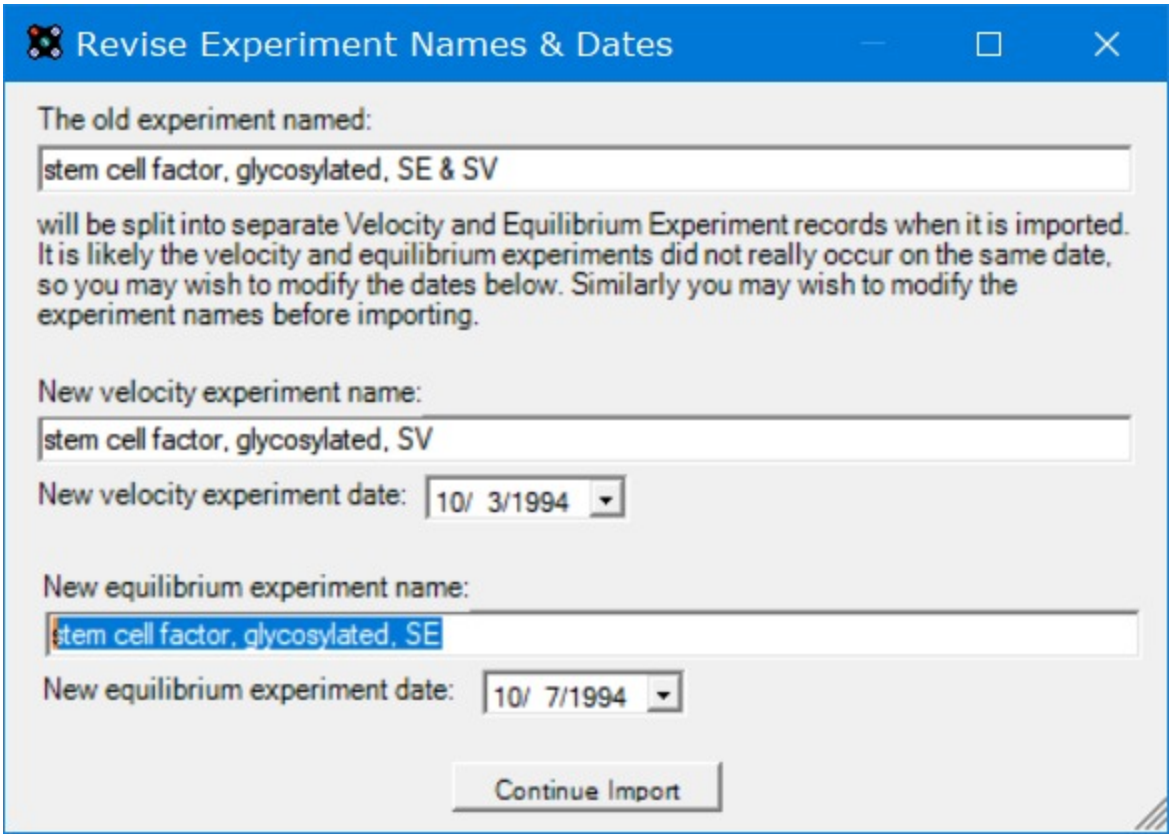
When the **Start import** button is pressed the form switches to the 'Import Log' tab to monitor and record the details about what data is (and is not) being imported. Records that are being imported will be listed in bold typeface, while those not being imported will be listed in normal typeface.



The image above shows a portion of the log from importing a version 1 database with only a few added samples. As the import proceeds the log is usually generated much too fast to read, but after it is complete you can (and should) scroll through from the beginning to confirm that everything went as expected. The log is a formatted rich-text document that you can print or copy to the Windows Clipboard using the provided buttons.

If in your old database you have added any density or viscosity data for new buffer components (solutes), you will be asked to provide the molar mass for the solute, and also the liquid density for any new solutes that are liquids. (This is information not present in version 1 databases, but which version 3 needs in order to convert between different concentration units.)

One major difference between version 1 and version 3 is that version 3 uses separate records (different tables) for sedimentation velocity and sedimentation equilibrium experiment information, whereas in version 1 both velocity and equilibrium results for the same sample were combined into a single record. That single record unfortunately assigned the same experiment name and same experiment date to both the velocity and equilibrium experiments. Therefore when the import process finds an Experiment record that contains both velocity and equilibrium information, the process will pause and display a dialog box like that shown below to allow you to confirm or change the name and experiment date for the new separated Velocity and Equilibrium Experiment records.



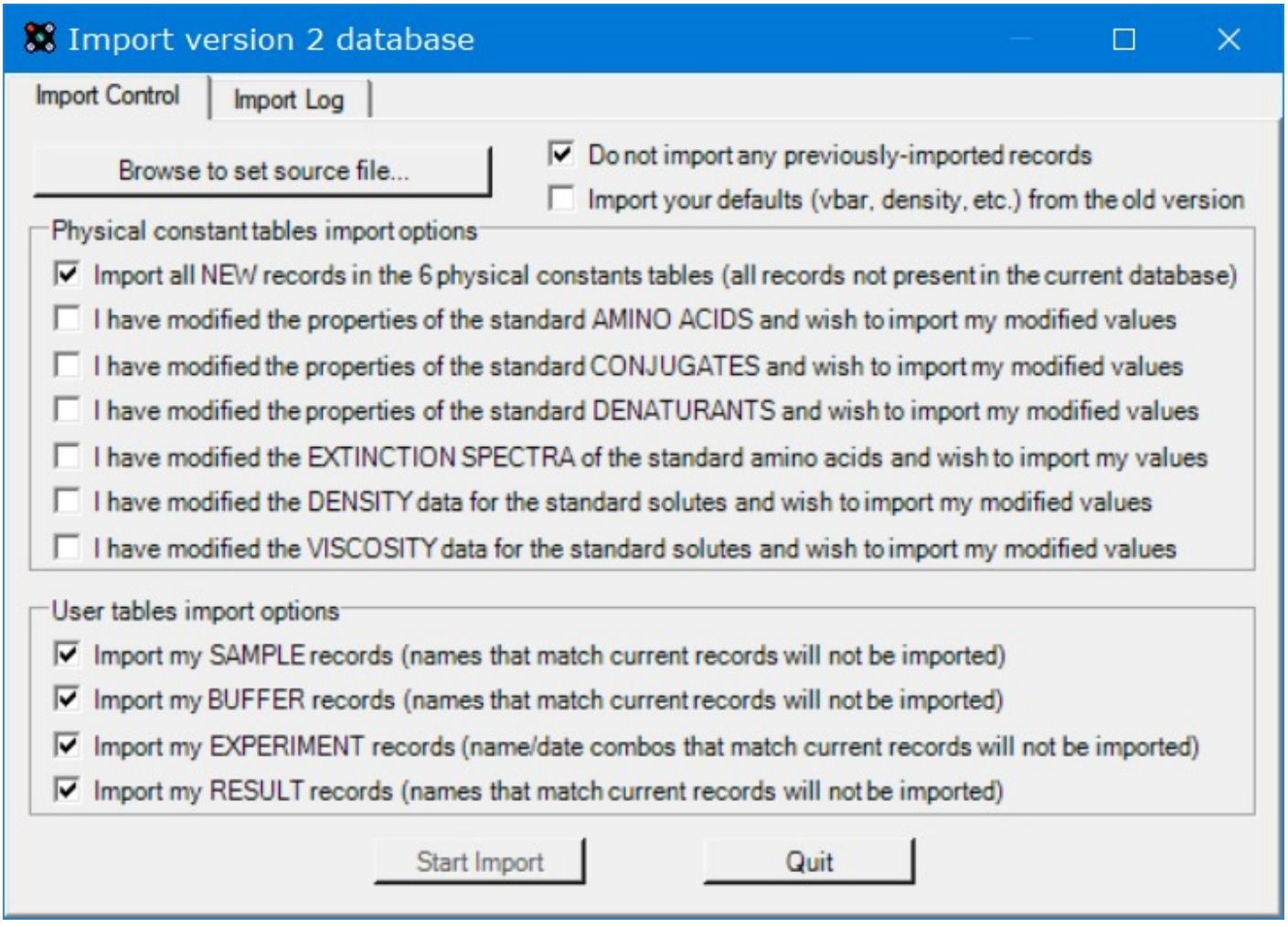
Please! If you encounter any errors during the import process, copy the import log into a WordPad or WORD document, and e-mail both that document and copies of the sednterp.mdb and phyconst.mdb files you were importing to John Philo (jphilo@mailway.com) so he can try to find and fix the problem for everyone.

Importing a database from SEDNTERP version 2 (beta)

SEDNTERP version 2 used a single SQLite database file (usually the file sednterp.db). SEDNTERP version 3 also uses a single SQLite database file, but it uses different table structures, additional tables, and some items (columns) within the 'user' portion have been moved to different tables. There are also significant differences between versions 2 and 3 regarding linking together Sample, Buffer, Experiment, and Hydrodynamic Result records.

Unfortunately the version 2 database did not store certain important information, and in some cases that means you will be asked to supply additional information as the import process proceeds (more details below).

Invoking the **Import version 2 database...** command from the File Menu will bring up this form:



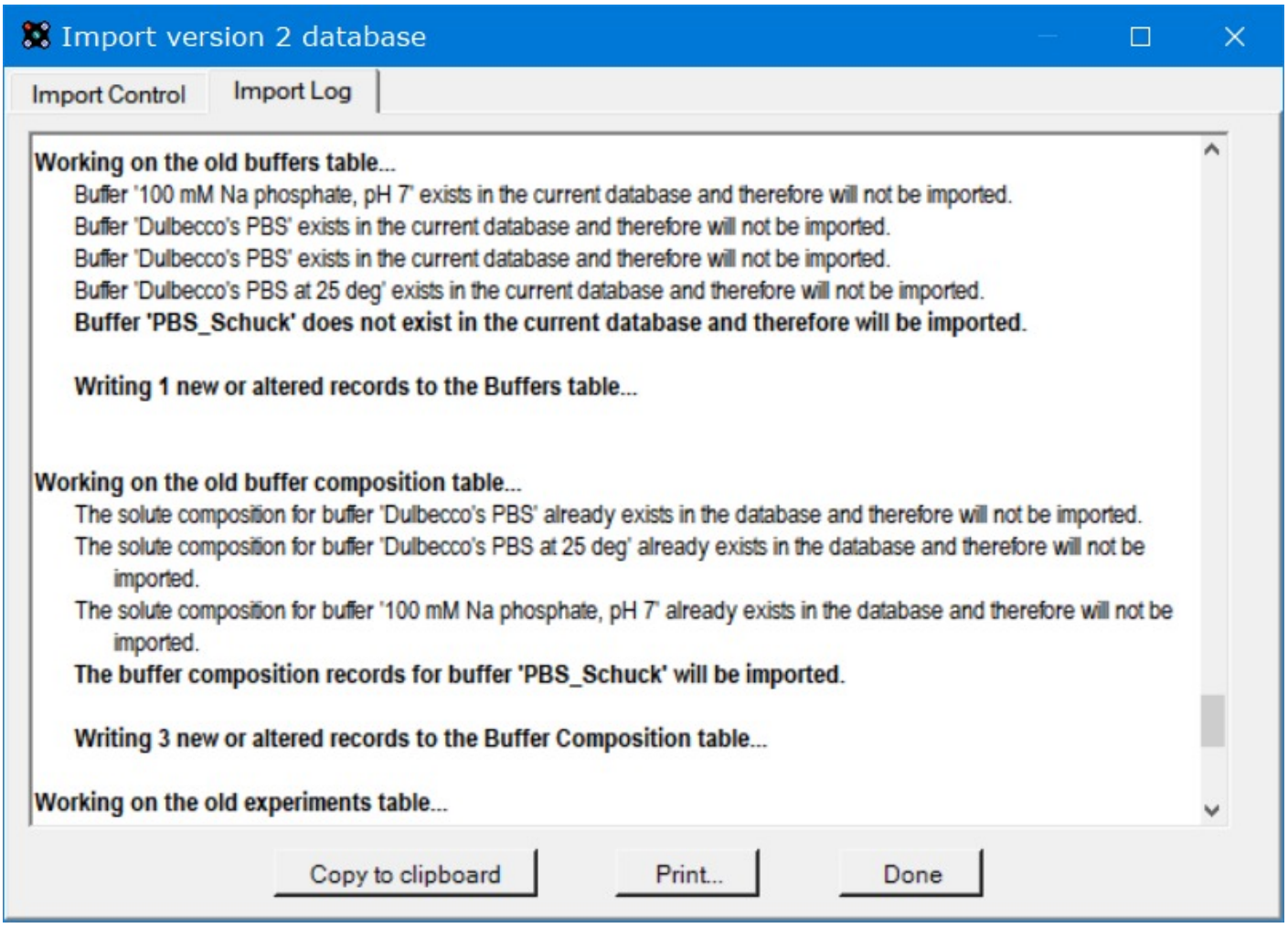
The **Browse to set source file...** button must be used to designate the folder where the version 2 database is located and its filename (usually sednterp.db). Once the file location is set the **Start import** button becomes active.

The options for controlling the import of the physical constants and user tables, as well as your old defaults, were already discussed above, and should be reviewed before you click the **Start import** button.

The sole remaining option is the 'Do not import any previously-imported records' check box at top right, which is on by default. As SEDNTERP 3 imports your old databases it marks the records it imports. Therefore if for some reason you wish to go through the import process again, ordinarily it would skip over those previously-imported records, but you can override that behavior by un-checking this box.

The **Quit** button will simply terminate the import, close this import control form, and return to the Main Data Entry form.

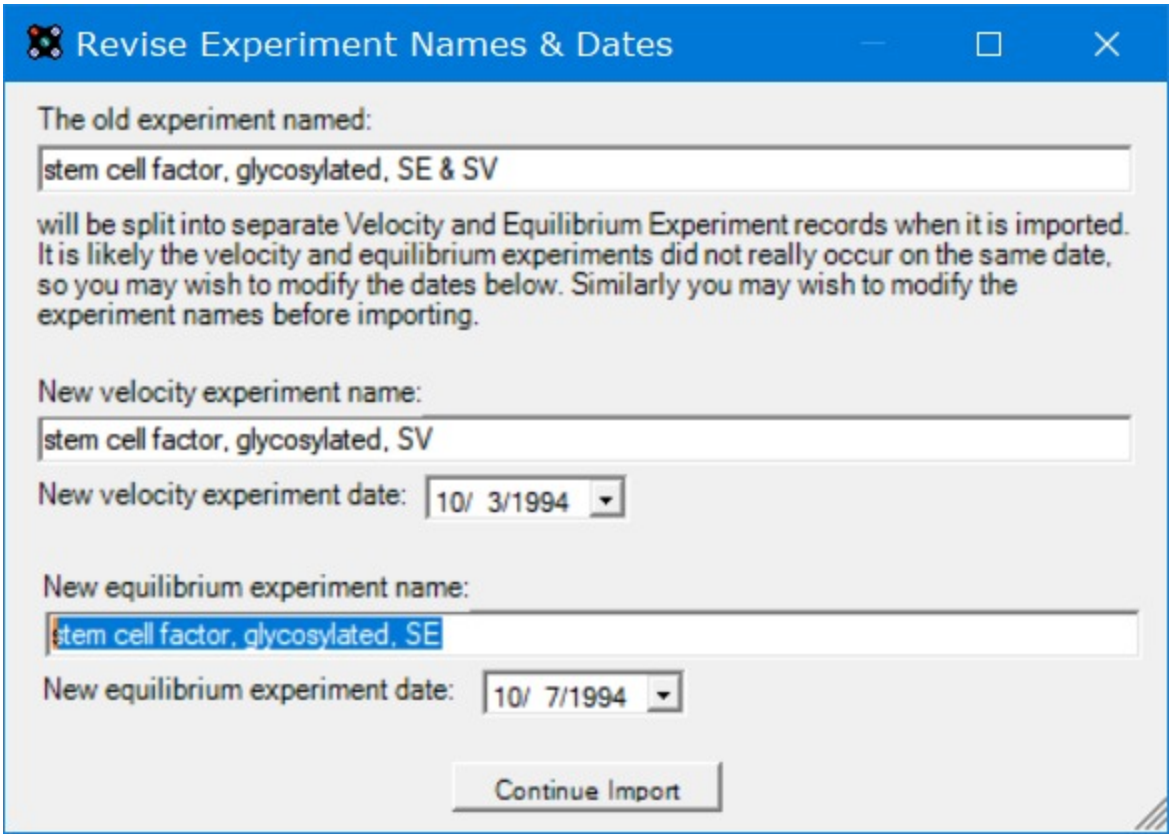
When the **Start import** button is pressed the form switches to the 'Import Log' tab to monitor and record the details about what data is (or is not) being imported. Records that are being imported will be listed in bold typeface, while those not being imported will be listed in normal typeface.



The image above shows a portion of the log from importing a version 2 database with only a few added samples and 1 added buffer. As the import proceeds the log is usually generated much too fast to read, but after it is complete you can (and should) scroll through from the beginning to confirm that everything went as expected. The log is a formatted rich-text document that you can print or copy to the Windows Clipboard using the provided buttons.

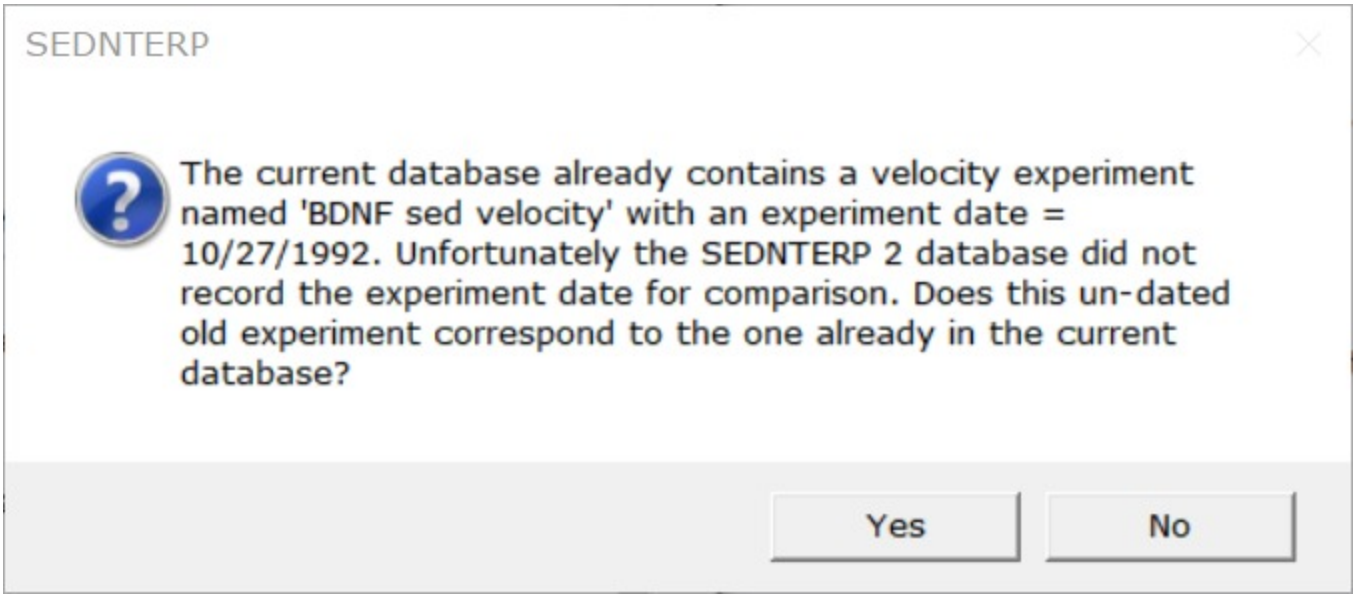
If you have added any density or viscosity data for new buffer components (solutes) to your old database, you will be asked to provide the molar mass for the solute, and also the liquid density for any new solutes that are liquids. (This is information that is usually not present in the version 2 database, but which version 3 needs in order to convert between different concentration units.)

One major difference between version 2 and version 3 is that version 3 uses separate records (different tables) for sedimentation velocity and sedimentation equilibrium experiment information, whereas in version 2 both velocity and equilibrium results for the same sample were combined into a single record. That single record unfortunately assigned the same experiment name to both the velocity and equilibrium experiments, and did not record an experiment date for either of these experiments. Therefore when the import process finds an Experiment record that contains both velocity and equilibrium information, the process will pause and display a dialog box like that shown below to allow you to confirm or change the name, and also to supply an actual experiment date, for the new separated Velocity and Equilibrium Experiment records.



Another problem with the version 2 databases is that version 2 allowed multiple Sample, Buffer, Experiment, or Result records to have the same name, even though the names were supposed to be unique for each record type (and they *must* be unique in version 3). Therefore if such duplicate names are found, only one record with each name will be imported.

As already mentioned, the version 2 database structure did not record the actual dates for experiments, and it also allowed multiple experiments with the name. This makes it difficult to tell whether an experiment in the old database truly matches one that has a matching name in the current (version 3) database, and thus whether it should or should not be imported. Therefore when there is ambiguity, the import will pause and you will be asked to confirm whether the un-dated old experiment corresponds to the one already in the current database, as illustrated below:



Lastly, it should be noted that the Sample records in the version 2 database did not record the sample's molar mass, and only stored its hydration if that was computed from composition. The import algorithm will attempt to find a molar mass for each sample in an Experiment record, but if that is not found, and if composition data for this sample is available, then it will be assumed that the molar mass should be calculated from the composition. If the imported record indicates that the hydration was not calculated from composition, it will then be set to the default hydration value.

Please! If you encounter any errors during the import process, copy the import log into a WordPad or WORD document, and e-mail both that document and copies of the sednterp.mdb and phyconst.mdb files you were importing to John Philo (jphilo@mailway.com) so he can try to find and fix the problem for everyone.

5 The Physical Properties Tables

5.1 The Amino Acids table

This table contains key data about each amino acid residue, such as its mass, partial specific volume, charge, molar refractive increment, and molar extinction coefficient at 280 nm. Typically there is no need for the user to alter or add to these data. However, this topic will provide details about the fields (column names) of this table in case you do need to modify them, and also will provide information about the source of these data.

Column descriptions:

ID - This first column should never be modified by the user! This value is used as a unique key to identify each amino acid record (each row in the table). If you are adding new rows to the table (such as a non-natural amino acid) the program will automatically fill in this entry with the next available value in the sequence.

Name - This field is the name of the amino acid (up to 25 characters), which must be different from all existing entries. This field cannot be empty *(NULL)*.

NameThree - This field is the three-letter code for this amino acid (which must be different from all existing entries). This field cannot be empty (*NULL)*.

NameOne - This field is the one-letter code for this amino acid. If entered, the code must be different from all existing entries, but this column can be *NULL* if you do not wish to assign a code (like the entry for disulfide). (Note that the database considers each *NULL* entry to be unique.)

Mr - This entry is the residue molar mass (not the mass of the free acid), for an average isotopic composition (not the monoisotopic mass). An entry in this column is required (it cannot be *NULL*). The values in the as-supplied database are the ones used by the Mascot mass spectroscopy software (Matrix Science), which are accurate to 4 decimal places (see table here).

Vbar - This entry is the residue partial specific volume (in mL/g, at 25 °C). An entry in this column is required (it cannot be *NULL*). The values in the as-supplied database are those from Table 2a of Ref. 59 (values originally from Cohn & Edsall, Ref. 4, except for cysteine, which comes from Zamyatnin, Ref. 69).

HipH7 - This entry is the hydration (moles of bound water per residue) for pH values above 4. No entry is required, and the entry is a real number (it does not have to be an integer). The values in the as-supplied database are from Kuntz (Ref. 51).

HipH4 - This entry is the hydration (moles of bound water per residue) for pH values of 4 or

below. No entry is required, and the entry is a real number (it does not have to be an integer). The values in the as-supplied database are from Kuntz (Ref. 51).

Extinction280 - This is the residue molar extinction coefficient at 280 nm. No entry is required. The values in the as-supplied database are from Edelhoch (Ref. 63).

Charge - This entry is the electric charge per residue, and is used in conjunction with the ***pK$_a$*** entry. If the pK_a entry is any negative value, the charge is considered permanent (non-titrable). If the pK_a entry is positive, the net charge will be calculated for the current pH based on that pK_a value, as detailed here. No entry is required.

pKa - This entry is the apparent *pK* of the charged group, and is used in conjunction with the **Charge** entry. If the pK_a entry is any negative value, the charge is considered permanent (non-titrable). If the pK_a entry is positive, the net charge will be calculated for the current pH based on that pK_a value, as detailed here. No entry is required. The residue pK_a values in the as-supplied database are from Segel (Ref. 66).

Refractivity - This entry is the molar refractivity for the residue, at 25 °C and a wavelength of 589.26 nm. No entry is required (but without any value the calculated refractive increment for the protein will not be accurate). The values in the as-supplied database are from Table 1 of Zhao *et al.* (Ref. 70), except that their value given for cysteine is actually that for cystine.

HasSpectum - This entry is an integer to indicate whether this conjugate has an associated UV absorbance spectrum in the database (in the Extinction table). The integer is used like a Boolean value: an entry of 0 is 'false' and indicates the conjugate is not a carbohydrate, and any other value is 'true' and indicates this conjugate should be treated as a carbohydrate. An entry is required (it cannot be *NULL*), but the database will automatically supply a default value of 0.

Modified - This field gives the date and time this record (row) was last modified. This column cannot be modified by the user.

5.2 The Conjugates table

This table contains key data about each defined conjugate, such as its mass, partial specific volume, charge, molar refractive increment, and molar extinction coefficient at 280 nm. Typically there is no need for the user to alter or add to these data. However, this topic will provide details about the fields (column names) of this table in case you do need to modify them, and also will provide information about the source of some of these data.

Column descriptions:

ID - This first column should never be modified by the user! This value is used as a unique key to identify each conjugate record (each row in the table). If you are adding new rows to the table the program will automatically fill in this entry with the next available value in the sequence.

Name - This field is the name of the conjugate (up to 30 characters), which must be different from all existing entries. This field cannot be empty *(NULL)*.

Mr - This entry is the conjugate molar mass, for an average isotopic composition (not the monoisotopic mass). An entry in this column is required (it cannot be *NULL*).

Vbar - This entry is the conjugate's partial specific volume (in mL/g, at 25 °C). An entry in this column is required (it cannot be *NULL*). The values in the as-supplied database are those from Tables 2b-2g of Ref. 59 (values excerpted from Durchschlag, Ref. 6).

HipH7 - This entry is the hydration (moles of bound water per conjugate) for pH values above 4. No entry is required, and the entry is a real number (it does not have to be an integer). The value for the new 10 kDa polyethylene glycol conjugate entry is based on data for 2 kDa PEG in ref. 92.

HipH4 - This entry is the hydration (moles of bound water per conjugate) for pH values of 4 or below. No entry is required, and the entry is a real number (it does not have to be an integer). The value for the new 10 kDa polyethylene glycol conjugate entry is based on data for 2 kDa PEG in ref. 92.

Extinction280 - This is the conjugate molar extinction coefficient at 280 nm. No entry is required. None of the conjugates in the as-supplied database have non-zero values.

Charge - This entry is the electric charge per conjugate, and is used in conjunction with the ***pK_a*** entry. If the pK_a entry is any negative value, the charge is considered permanent (non-titrable). If the pK_a entry is positive, the net charge will be calculated for the current pH based on that pK_a value, as detailed here. No entry is required.

pKa - This entry is the apparent *pK* of the charged group, and is used in conjunction with the **Charge** entry. If the pKa entry is any negative value, the charge is considered permanent (non-titratable). If the pK_a entry is positive, the net charge will be calculated for the current pH based on that pK_a value, as detailed here. No entry is required.

Carbo - This entry is an integer used to indicate whether this conjugate should be treated as a carbohydrate for purposes of calculating the partial specific volume for proteins denatured in 6 M guanidine hydrochloride (and therefore binds an additional denaturant molecule per mole). See Calculating denatured vbar estimates. The integer is used like a Boolean value: an entry of 0 is 'false' and indicates the conjugate is not a carbohydrate, and any other value is 'true' and indicates this conjugate should be treated as a carbohydrate. An entry is required, but the database will automatically supply a default value of 0.

Refractivity - This entry is the molar refractivity for the conjugate, at 25 °C and a wavelength of 589.26 nm. No entry is required (but without any value the calculated refractive increment for the conjugated protein will not be accurate). The values in the as-supplied database are from various sources. If a refractivity could not be found in the literature, but a *dn/dc* value was found in the literature, a refractivity was back-calculated to give that reported *dn/dc*. If neither the refractivity or *dn/dc* value could be located, the predicted refractivity from ChemSpider (calculated by the ACD/Labs Percepta Platform) was used (when available).

HasSpectum - This entry is an integer to indicate whether this conjugate has an associated UV absorbance spectrum in the database (in the Extinction table). The integer is used like a Boolean value: an entry of 0 is 'false' and indicates the conjugate is does not have a stored extinction spectrum, and any other value is 'true' and indicates this conjugate does have a stored extinction spectrum. An entry is required (it cannot be *NULL*), but the database will automatically supply a default value of 0. None of the conjugates defined in the as-supplied database have associated spectra.

Modified - This field gives the date and time this record (row) was last modified. This column cannot be modified by the user.

5.3 The Solutes table

For each defined solute this Solutes table contains pointers to the correct record within the density, viscosity, and refractive index data tables, as well as some parameters needed to convert molar concentrations to molal, % by weight, or % by volume concentration scales.

It is important to note that if you are going to add a new solute, you must first add its density, viscosity, and/or refractive index data into those tables, and then create the corresponding new record in this Solutes table.

Column descriptions:

ID - This first column should never be modified by the user! This value is used as a unique key to identify each record (each row in the table). If you are adding new rows to the table (to create a new solute) the database will automatically fill in this entry with the next available value in the sequence.

Name - This field is the name of the solute (up to 40 characters), which must be different from all existing entries. The same name should be used for the corresponding entries in the density, viscosity, and refractive index data tables. This field cannot be empty (null).

SoluteRhoID - This column holds the ID number (the number in the ID column) for the row in the

solute Density data table which holds the density coefficients for this solute. If no density data exists for this solute, this column should be set as *NULL* (use CTRL+0 to enter *NULL*).

SoluteViscD - This column holds the ID number (the number in the ID column) for the row in the solute Viscosity data table which holds the density coefficients for this solute. If no density data exists for this solute, this column should be set as *NULL* (use CTRL+0 to enter *NULL*).

SoluteRefrIndexID - This column holds the ID number (the number in the ID column) for the row in the solute Refractive Index data table which holds the density coefficients for this solute. If no density data exists for this solute, this column should be set as *NULL* (use CTRL+0 to enter *NULL*).

LiquidDensity - This column should contain the density (g/mL) at 20 °C for solutes which are liquids. That value is required in order to be able to enter concentrations on a % by volume scale and convert them to the molar scale used internally. For solutes that are not liquids this column should be set as *NULL* (use CTRL+0 to enter *NULL*).

SoluteMW - This entry is the molar mass of the solute. This value is required to convert concentrations on a molal or % by weight scale to the molar scale used internally.

Modified - This field gives the date and time this record (row) was last modified. This column cannot be modified by the user.

5.4 The Solute Density, Viscosity, and Refractive Index tables

These three tables contain the polynomial coefficients and other data used in interpolating the density, viscosity, and refractive index of a buffer based on its solute composition. These 3 tables will be discussed together because they all have virtually the same structure.

Note that the records in these 3 tables are logically linked to a 'master' table of defined solutes, the Solutes table. If you are going to add density, viscosity, or refractive index data for a new solute, you must first add all of those new entries, and then create the corresponding new record in the Solutes table.

Column descriptions:

ID - This first column should never be modified by the user! This value is used as a unique key to identify each record (each row in the table). If you are adding new rows to the table (to create a new solute, or to add information that is missing about a current solute) the database will automatically fill in this entry with the next available value in the sequence.

Name - This field is the name of the solute (up to 40 characters), which must be different from all existing entries, and which must match the name used in the Solutes table. This field cannot be empty *(NULL)*.

a, b, c, d, e, f - These columns contain the coefficients for the fitting function used to compute the density, viscosity, or refractive index. Note that the viscosity table does not have a column for the 'f' parameter. Many solutes do not require all six of these parameters for a sufficiently accurate fit; unused parameters can be either explicitly entered as zero or left as *NULL*.

Cmax - This entry gives the maximum molar concentration covered by the data used to fit the polynomial coefficients (that is, the maximum concentration where we know the fitting function is sufficiently accurate). An entry is not absolutely required, but without this entry the program cannot signal that the user has entered a concentration that is out of range.

ErrorMax - This entry gives the maximum error (multiplied by 10,000) between the fitting function and the experimental data over the concentration range up to C_{max}. This entry is optional and intended only as informational; it's value is not used in any calculations.

Source - This text entry (80 characters max) is used to indicate the source of the experimental values (a reference, person, or lab).

Modified - This field gives the date and time this record (row) was last modified. This column cannot be modified by the user.

5.5 The Defined Denaturants table

This table contains physical parameter values that are used in calculating (estimating) the partial specific volume of denatured proteins. As supplied, the database contains only two defined denaturants, 8 M urea and 6 M guanidine hydrochloride. Those are the only two defined conditions because to the authors' knowledge those are the only ones where the necessary experimental data have been published. Although it is therefore unlikely the user will have the necessary data to create other defined denaturants, or wish to modify the as-supplied values, the entries in this table will be described below in case changes or additions are needed.

Column descriptions:

ID - This first column should never be modified by the user! This value is used as a unique key to identify each amino acid record (each row in the table). If you are adding new rows to the table the program will automatically fill in this entry with the next available value in the sequence.

Name - This field is the name of the defined denaturant (up to 25 characters), which must be

different from all existing entries. This field cannot be empty (*NULL*).

RhoMin - This field contains the density of the defined denaturant solution (g/mL) at 20 ^{o}C.

v3 - This field contains the partial specific volume of the denaturant.

g3 - This field contains the grams of denaturant per gram of water in the denaturant solution.

Md - This field contains the molar mass of the denaturant (Da).

Carbo - This entry is an integer used to indicate whether the calculations should assume that this denaturant also binds to conjugates marked as being carbohydrates (which is true for guanidine hydrochloride). See Calculating denatured vbar estimates for more details about how this is handled. The integer is used like a Boolean value: an entry of 0 is 'false' and indicates the denaturant does not bind to carbohydrates, and any other value is 'true' and indicates no binding to carbohydrates should be included in the calculations. An entry is required, but the database will automatically supply a default value of 0.

Modified - This field gives the date and time this record (row) was last modified. This column cannot be modified by the user.

5.6 The Extinction table

This table contains values of molar extinction *vs.* wavelength for tryptophan, tyrosine, phenylalanine, and disulfide (plus any other conjugates or non-natural amino acids the user many have added). These data are used to calculate and display the extinction spectra shown on the sample composition form.

Since each record (row) in the table contains only one (wavelength, extinction) pair, defining the spectrum for each compound requires creating many rows (36 rows for the normal range from 250 to 320 nm with a 2 nm point interval). Consequently, if you wish to add a new compound to the table, after you push the **Add new row button** on the Edit Physical Properties Tables form the program will ask you what range of wavelengths and what wavelength interval (spacing) you wish to use, and it will then pre-fill the wavelength entries as you add new rows.

Note that the graphs displayed on the Sample Composition form will only include the wavelength range covered by **all** of the compounds in this table, and will never include wavelengths below 250 nm (because the peptide bond contributions in that region cannot be calculated accurately). You may use a data point interval larger or smaller than 2 nm for new compounds, however that will cause the data points on the graphs to be interpolated to a common interval.

Column descriptions:

ID - This first column should never be modified by the user! This value is used as a unique key to identify each record (each row in the table). If you are adding new rows to the table (to add a spectrum for a non-natural amino acid or a conjugate) the database will automatically fill in this entry with the next available value in the sequence.

Name - This field is the name of the amino acid or conjugate (up to 30 characters), which must match the name used in the Amino Acids or Conjugates tables. This field cannot be empty *(NULL)*.

nm - This field is the wavelength in nm. It is a real number (it does not have to be an integer). This field cannot be empty *(NULL)*.

Extinction - This field is the molar extinction at the wavelength in the column to the left. This field cannot be empty *(NULL)*.

Modified - This field gives the date and time this record (row) was last modified. This column cannot be modified by the user.

6 Bibliography

6.1 Reference 1

R.J. Goldberg, J. Phys. Chem., 1953, 57, 194.

6.2 Reference 2

Williams, J. W., van Holde, K. E., Baldwin, R. L., and Fujita, H. (1958). The theory of sedimentation analysis. *Chemical Reviews.* 58, 715-744.

6.3 Reference 3

E.F. Cassasa and H. Eisenberg Adv. in Prot. Chem., 1964, 19, 287.

6.4 Reference 4

E.J. Cohn and J.T. Edsall 'Proteins, Amino Acids and Peptides as Ions and Dipolar Ions', Rheinhold, New York, 1943, Chapter 4, p. 157.

6.5 Reference 5

T.L. McMeekin and K. Marshall, Science, 1952, 116, 142.

6.6 Reference 6

H. Durchschlag, 'Thermodynamic Data for Biochemistry and Biotechnology', H.-J. Hinz, ed., Springer-Verlag, New York, 1986, Chapter 3, p. 45.

6.7 Reference 7

D.W. Kupke, 'Physical Principles and Techniques of Protein Chemistry', S.J. Leach, ed., Academic Press, New York, 1973, Part C, Chapter 1, p. 1.

6.8 Reference 8

T. Svedberg 'The Ultracentrifuge', T. Svedberg and K.O. Pederson, eds., Johnson Reprint Corporation, New York, 1940.

6.9 Reference 9

H.K. Schachman, 'Methods in Enzymology Volume IV', S.P. Colowick and N.O. Kaplan, eds., Academic Press, New York, 1957, Chapter 2, p. 32.

6.10 Reference 10

H.K. Schachman, 'Ultracentrifugation in Biochemistry', Academic Press, New York, 1959.

6.11 Reference 11

C.H. Chervenka, 'A Manual of Methods for the Analytical Ultracentrifuge', Spinco Division of Beckman Instruments, 1970, Palo Alto, California.

6.12 Reference 12

R. Eason, 'Centrifugation (2nd Edition) a Practical Approach', D. Rickwood, ed., IRL Press, 1984, Chapter 8, p. 251.

6.13 Reference 13

G. Cohen and H. Eisenberg, Biopolymers, 1968, 6, 26.

6.14 Reference 14

Yphantis, D. A. (1964). Equilibrium ultracentrifugation of dilute solutions. *Biochemistry* 3, 297-317.

6.15 Reference 15

W.F. Harrington and G. Kegeles, 'Methods in Enzymology Volume XXVII', C.H.W. Hirs and S.N. Timasheff, eds., Academic Press, New York, 1973, Chapter 13, p. 306.

6.16 Reference 16

W.F. Harrington, Fractions, 1975, 1, 10.

6.17 Reference 17

H. Fujita, 'Foundations of Ultracentrifugal Analysis', J. Wiley and Sons, New York, 1975.

6.18 Reference 18

C.R. Cantor and P.R. Schimmel, 'Biophysical Chemistry, Part II', W.H. Freeman and Co., San Francisco, 1990, Chapter 10, p. 539.

6.19 Reference 19

C.R. Cantor and P.R. Schimmel, 'Biophysical Chemistry, Part II', W.H. Freeman and Co., San Francisco,

1980, Chapter 12, p. 643.

6.20 Reference 20

J.P. Johnston and A.G. Ogston, Trans. Faraday Soc., 1946, 42, 789.

6.21 Reference 21

J.M. Creeth and C.G. Knight, Biochem. Biophys. Acta, 1965, 102, 549.

6.22 Reference 22

C.R. Cantor and P.R. Schimmel, 'Biophysical Chemistry, Part II', W.H. Freeman and Co., San Francisco, 1980, Chapter 11, p. 591.

6.23 Reference 23

C.R. Cantor and P.R. Schimmel, 'Biophysical Chemistry, Part II', W.H. Freeman and Co., San Francisco, 1980 p. 613.

6.24 Reference 24

J.R. Cann, 'Interacting Macromolecules: The Theory and Practice of Their Electrophoresis, Ultracentrifugation and Chromatography', Academic Press, New York, 1970, Chapter III, p. 93.

6.25 Reference 25

D.A Yphantis and D.F. Waugh, J. Phys. Chem., 1956, 60, 623.

6.26 Reference 26

M.L. Johnson, J.C. Correia, D.A. Yphantis and H.R. Halvorson Biophysical J., 1981, 36, 575.

6.27 Reference 27

Teller, D.C. 'Methods in Enzymology Volume XXVII', C.H.W. Hirs and S.N. Timasheff, eds., Academic Press, New York, 1973, Chapter 14, p. 346.

6.28 Reference 28

Kumosinski, T. F. and Pessen, H. (1985). Structural interpretation of hydrodynamic measurements of proteins in solution through correlations with X-ray data. *Methods Enzymol.* 117, 154-182.

6.29 Reference 29

Oncley, J. L. (1941). Evidence from physical chemistry regarding the size and shape of protein molecules from ultracentrifugation, diffusion, viscosity, dielectric dispersion, and double refraction of flow. *Ann. NY Acad. Sci.* 41, 121-150.

6.30 Reference 30

Teller, D. C., Swanson, E., and de Haën, C. (1979). The translational friction coefficient of proteins. *Methods Enzymol.* 61, 104-124.

6.31 Reference 31

Scheraga, H. A. and Mandelkern, L. (1953). Consideration of the hydrodynamic properties of proteins. *Journal of the American Chemical Society* 75, 179-184.

6.32 Reference 32

Harding, S. E. (1997). The intrinsic viscosity of biological macromolecules. Progress in measurement, interpretation and application to structure in dilute solution. *Progress in biophysics and molecular biology* 68, 207-262.

Harding, S. E. (1987). A general method for modeling macromolecular shape in solution: A graphical (II-G) intersection procedure for triaxial ellipsoids. *Biophys. J.* 51, 673-680.

.

6.33 Reference 33

Teller, D. C. (1976). Accessible area, packing volumes and interaction surfaces of globular proteins. *Nature* 260, 729-731.

6.34 Reference 34

Höiland, H. (1986). Partial molar volumes of biochemical model compounds in aqueous solutions. In: Thermodynamic Data for Biochemistry and Biotechnology. Springer-Verlag, Berlin, pp. 17-44.

6.35 Reference 35

J.H.C. Steele, C. Tanford and J.A. Reynolds, 'Methods in Enzymology Volume 48', C.H.W. Hirs and S.N. Timasheff, eds., Academic Press, Orlando, Florida, 1978, Chapter 2, p. 18.

6.36 Reference 36

J.A. Reynolds and D.R. McCaslin, 'Methods in Enzymology Volume 117', C.H.W. Hirs and S.N. Timasheff, eds., Academic Press, Orlando, Florida, 1985, Chapter 3, p. 41.

6.37 Reference 37

Prakash, V. and Timasheff, S. N. (1985). Calculation of partial specific volumes of proteins in 8 M urea solution. Methods Enzymol. 117, 53-60.

6.38 Reference 38

T. Arakawa and S.N. Timasheff, 'Methods in Enzymology Volume 117', C.H.W. Hirs and S.N. Timasheff, eds., Academic Press, Orlando, Florida, 1985, Chapter 5, p. 60.

6.39 Reference 39

D.W. Kupke, 'Physical Principles and Techniques of Protein Chemistry, part C', S.J. Leach, ed., Academic Press, New York, 1973, Chapter 1, p. 1.

6.40 Reference 40

Perkins, S. J. (1986). Protein volumes and hydration effects. The calculations of partial specific volumes, neutron scattering matchpoints and 280-nm absorption coefficients for proteins and glycoproteins from amino acid sequences. *Eur. J. Biochem.* 157, 169-180.

6.41 Reference 41

Gibbons, R. A. (1972). *Glycoproteins: Their Composition, Structure and Function, Part 1*. A. Gottschalk, ed. Elseviier, Amsterdam, pp. 31-140.

6.42 Reference 42

Lee, J. C. and Timasheff, S. N. (1974). Partial specific volumes and interactions with solvent components of proteins in guanidine hydrochloride. *Biochemistry* 13, 257-265.

6.43 Reference 43

Lee, J. C., Gekko, K., and Timasheff, S. N. (1979). Measurements of preferential solvent interactions by densimetric techniques. *Methods Enzymol.* 61, 26

6.44 Reference 44

R.C. Weast (editor), 'CRC Handbook of Chemistry and Physics, 67th Edition', CRC Press, Inc., Boca Raton, Florida, 1986, p. F-5.

6.45 Reference 45

F. Steckel and S. Szapiro, Trans. Faraday Soc., 1963, 59, 331.

6.46 Reference 46

M.L. Johnson and S.G. Frazsier, 'Methods in Enzymology Volume 117', C.H.W. Hirs and S.N. Timasheff, eds., Academic Press, Orlando, Florida, 1985, Chapter 16, p. 301.

6.47 Reference 47

J.T. Yang, 'Recent Advances in Protein Chemistry Volume 16', C.B. Anfinsen, M.L Anson, K. Bailey, and J.T. Edsall, eds., Academic Press, New York, 1961, p. 323.

6.48 Reference 48

J.F. Johnson, J.R. Martin and R.S. Porter, 'Physical Methods of Chemistry', A. Weissberger and B.W. Rossiter, eds., Vol. 1, Part VI, Wiley, New York, 1977, p. 63.

6.49 Reference 49

R.C. Weast (editor), 'CRC Handbook of Chemistry and Physics, 67th Edition', CRC Press, Inc., Boca Raton, Florida, 1986, p. F-37.

6.50 Reference 50

K.E. van Holde, 'Physical Biochemistry 2nd Edition', Prentice-Hall, New Jersey, 1985, p. 117.

6.51 Reference 51

Kuntz, I. D. (1971). Hydration of macromolecules. IV. Polypeptide conformation in frozen solutions. *J Am. Chem. Soc.* 93, 516-518.

6.52 Reference 52

Lüscher-Mattli, M. (1986). Thermodynamic parameters of biopolymer-water systems. In: *Thermodynamic data for biochemistry and biotechnology*. Springer-Verlag, Berlin, pp. 276-294.

6.53 Reference 53

Mrevlishvili, G. M. (1986). Heat capacities of biological macromolecules. In: *Thermodynamic Data for Biochemistry and Biotechnology*. AnonymousSpringer-Verlag, Berlin, pp. 148-176.

6.54 Reference 54

Kuntz, I. D. and Kauzmann, W. (1974). Hydration of proteins and polypeptides. *Adv. Protein Chem.* 28, 239-345.

6.55 Reference 55

Pessen, H. and Kumosinski, T. F. (1985). Measurements of protein hydration by various techniques. *Methods Enzymol.* 117, 219-255.

6.56 Reference 56

E.G. Richards and H.K. Schachman, J. Amer. Chem. Soc., 1957, 79, 5324.

6.57 Reference 57

M.W. Kirschner and H.K. Schachman, Biochemistry, 1971, 10, 1900.

6.58 Reference 58

J. G. de la Torre, "Sedimentation Coefficiients of Complex Biological Particles", in *Analytical Ultracentrifugation in Biochemistry and Polymer Science*, Royal Society of Chemistry (Edited by S. Harding and A. Rowe), 1992, Chapter 18, pp. 333-345.

6.59 Reference 59

T.M. Laue, B.D. Shah, T.M. Ridgeway and S.L. Pelletier, Analytical Ultracentrifugation in Biochemistry and Polymer Science, Royal Society of Chemistry (Edited by S. Harding and A. Rowe), 1992, pp. 90-125.

6.60 Reference 60

W.F. Stafford, Analytical Ultracentrifugation in Biochemistry and Polymer Science, Royal Society of Chemistry (Edited by S. Harding and A. Rowe), 1992, pp. 359-393.

6.61 Reference 61

Lawrence J. Henderson (1908). "Concerning the relationship between the strength of acids and their capacity to preserve neutrality". *Am. J. Physiol*. 21: 173–179.

6.62 Reference 62

Hasselbalch, K. A. (1917). The calculation of the hydrogen content in blood from free and combined carbonic acid, and the oxygen compound of the blood as the function of the hydrogen content.

Biochemica Zeitschrift 78, 112-144.

6.63 Reference 63

Edelhoch, H. (1967). Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry* 6, 1948-1954.

6.64 Reference 64

Gill, S. C. and Von Hippel, P. H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. *Analytical. Biochemistry* 182, 319-326.

6.65 Reference 65

Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995). How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* 4, 2411-2423.

6.66 Reference 66

Segel, I. H. (1975). Biochemical calculations. Wiley, New York.

6.67 Reference 67

Fasman, G. D. (1989). Practical handbook of biochemistry and molecular biology. CRC press.

6.68 Reference 68

Svedberg, T. and Pedersen, K. O. (1940). The Ultracentrifuge. Clarendon Press, Oxford

6.69 Reference 69

Zamyatnin, A. A. (1972). Protein volume in solution. *Prog. Biophys. Mol. Biol.* 24, 107-123.

6.70 Reference 70

Zhao, H., Brown, P. H., and Schuck, P. (2011). On the distribution of protein refractive index increments. *Biophys. J.* 100, 2309-2317.

6.71 Reference 71

Batchelor, G. K. (1972). Sedimentation in a dilute dispersion of spheres. *Journal of fluid mechanics.* 52, 245-268.

6.72 Reference 72

Brady, J. F. and Durlofsky, L. J. (1988). The sedimentation rate of disordered suspensions. *Phys. Fluids* 31, 717-727.

6.73 Reference 73

Harding, S. E. and Johnson, P. (1985). The concentration-dependence of macromolecular parameters. *Biochemical. Journal* 231, 543-547.

6.74 Reference 74

Scott, D. J., Harding, S. E., and Winzor, D. J. (2014). Concentration dependence of translational diffusion coefficients for globular proteins. *Analyst* 139, 6242-6248.

6.75 Reference 75

Kingsbury, J. S., Saini, A., Auclair, S. M., Fu, L., Lantz, M. M., Halloran, K. T., Calero-Rubio, C., Schwenger, W., Airiau, C. Y., Zhang, J., and Gokarn, Y. (2020). A single molecular descriptor to predict solution behavior of therapeutic antibodies. *Science Advances*. 6, eabb0372 (prepub ahead of print)

6.76 Reference 76

Pátek, J., Hrubý, J., Klomfar, J., Soucková, M. and Harvey, A.H. (2009). Reference correlations for thermophysical properties of liquid water at 0.1 MPa. *Journal of Physical and Chemical Reference Data*, 38(1), pp. 21-29.

6.77 Reference 77

Rukes, B. and Dooley, R. B. (2020). Guideline on the Use of Fundamental Physical Constants and Basic Constants of Water, IAPWS G5-01 (2020). *International Association for the Properties of Water and Steam, Gaithersburg, Maryland, USA,* 1-7.

6.78 Reference 78

Philo, J. S. (1997). An improved function for fitting sedimentation velocity data for low-molecular-weight solutes. Biophys. J. 72, 435-444.

6.79 Reference 79

P. Schuck (2000). Size distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophysical Journal* 78, 1606-1619.

6.80 Reference 80

Lewis, M. S. and Junghans, R. P. (2000). Ultracentrifugal analysis of molecular mass of glycoproteins of unknown or ill-defined carbohydrate composition. *Methods Enzymol.* 321, 136-149.

6.81 Reference 81

Lu, Y., Harding, S. E., Turner, A., Smith, B., Athwal, D. S., Grossmann, J. G., Davis, K. G., and Rowe, A. J. (2008). Effect of PEGylation on the solution conformation of antibody fragments. *J. Pharm. Sci.* 97, 2062-2079.

6.82 Reference 82

Cho, C. H., Urquidi, J., Singh, S., and Robinson, G. W. (1999). Thermal offset viscosities of liquid H2O, D2O, and T2O. *The. Journal of Physical. Chemistry. B* 103, 1991-1994.

6.83 Reference 83

Harvey, A. H., Gallagher, J. S., and Sengers, J. M. H. L. (1998). Revised formulation for the refractive index of water and steam as a function of wavelength, temperature and density. *Journal of Physical. and Chemical. Reference. Data* 27, 761-774.

6.84 Reference 84

Odhner, H. and Jacobs, D. T. (2012). Refractive index of liquid D2O for visible wavelengths. *Journal of Chemical. & Engineering. Data* 57, 166-168.

6.85 Reference 85

Hansen, S. (2004). Translational friction coefficients for cylinders of arbitrary axial ratios estimated by Monte Carlo simulation. *The Journal of chemical physics.* 121, 9111-9115.

6.86 Reference 86

Weiner, O. (1910). Zur theorie der refraktionskonstanten berichte über die verhändlungen der königlich sächsischen gesellschaft der wissen-schaften zu Leipzig. *Mathematisch-Physikalische Klasse.* 62, 256-268.

6.87 Reference 87

Perlmann, G. E. and Longsworth, L. G. (1948). The specific refractive increment of some purified proteins. *J. Am. Chem. Soc.* 70, 2719-2724.

6.88 Reference 88

McMeekin, T. L., Groves, M. L., and Hipp, N. J. (1964). Refractive indices of amino acids, proteins, and related substances. In: *Amino Acids and Serum Proteins*. J. Stekol, ed. ACS Publications, Washington, D.C.,

6.89 Reference 89

Tirosh, O., Barenholz, Y., Katzhendler, J., and Priev, A. (1998). Hydration of polyethylene glycol-grafted liposomes. *Biophysical. journal* 74, 1371-1379.

6.90 Reference 90

Creeth, J. M. and Knight, C. G. (1965). On the estimation of the shape of macromolecules from sedimentation and viscosity measurements. *Biochim. Biophys. Acta* 102, 549-558.

6.91 Reference 91

Kinell, P. O. (1959). On the determination of molecular weight averages from sedimentation and diffusion data. *Arkiv für Kemi* 31, 327-336.

6.92 Reference 92

Tirosh, O., Barenholz, Y., Katzhendler, J., and Priev, A. (1998). Hydration of polyethylene glycol-grafted liposomes. *Biophysical Journal* 74, 1371-1379.

6.93 Reference 93

Thurlkill, R. L., Grimsley, G. R., Scholtz, J. M., and Pace, C. N. (2006). pK values of the ionizable groups of proteins. *Protein Sci.* 15, 1214-1218.

6.94 Reference 94

Durchschlag, H. and Zipper, P. (2001). Comparative investigations of biopolymer hydration by physicochemical and modeling techniques. *Biophys. Chem.* 93, 141-157.

6.95 Reference 95

Traube, J. (1899). *Samml. Chem. Vortr.* 4, 255-332.

6.96 Reference 96

Wolf, D. and Kudish, A. I. (1975). Absolute viscosity of deuterium oxide (oxygen-18) between 15 and 35 degrees. *The Journal of Physical Chemistry* 79, 1481-1482.

6.97 Reference 97

Philo, J. S. (2023). SEDNTERP: a calculation and database utility to aid interpretation of analytical ultracentrifugation and light scattering data. *Eur. Biophys. J.* 52, 233-266.

6.98 Reference 98

Wen, J., Arakawa, T., and Philo, J. S. (1996). Size-exclusion chromatography with on-line light-scattering, absorbance, and refractive index detectors for studying proteins and their interactions. *Anal. Biochem.* 240, 155-166.

6.99 Reference 99

Kendrick, B. S., Kerwin, B. A., Chang, B. S., and Philo, J. S. (2001). Online size-exclusion high-performance liquid chromatography light scattering and differential refractometry methods to determine degree of polymer conjugation to proteins and protein-protein or protein-ligand association states. *Anal. Biochem.* 299, 136-146.

6.100 Reference 100

Arakawa, T., Ejima, D., Li, T., and Philo, J. S. (2010). The critical role of mobile phase composition in size exclusion chromatography of protein pharmaceuticals. *J Pharm. Sci* 99, 1674-1692.

6.101 Reference 101

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (1992). Numerical recipes in FORTRAN 77: the art of scientific computation. Cambridge University Press, Cambridge, England.

6.102 Reference 102

Fleming, P. J., Correia, J. J., and Fleming, K. G. (2023). Revisiting macromolecular hydration with HullRadSAS. *Eur. Biophys. J.* 52, 215-224.

6.103 Reference 103

Harding, S. E. (1997). The intrinsic viscosity of biological macromolecules. Progress in measurement, interpretation and application to structure in dilute solution. *Prog. Biophys. Mol. Biol.* 68, 207

6.104 Reference 104

Fleming, P. J., Correia, J. J., and Fleming, K. G. (2024). The molecular basis for hydrodynamic properties of PEGylated human serum albumin. *Biophys. J.* 123, 2379-2391.

6.105 Reference 105

Svergun, D. I., Richard, S., Koch, M. H., Sayers, Z., Kuprin, S., and Zaccai, G. (1998). Protein hydration in solution: experimental observation by x-ray and neutron scattering. *Proc. Natl. Acad. Sci. Usa* 95, 2267-2272.

6.106 Reference 106

Seki, Y., Tomizawa, T., Khechinashvili, N. N., and Soda, K. (2002). Contribution of solvent water to the solution X-ray scattering profile of proteins. *Biophys. Chem.* 95, 235-252.

6.107 Reference 107

Dill, K. A. and Bromberg, S. (2010). Molecular driving forces: statistical thermodynamics in biology, chemistry, physics, and nanoscience. Garland Science, New York, pp. 327

6.108 Reference 108

Wright, R. T., Hayes, D., Sherwood, P. J., Stafford, W. F., and Correia, J. J. (2018). AUC measurements of diffusion coefficients of monoclonal antibodies in the presence of human serum proteins. *Eur. Biophys. J.* 47, 709-722.

6.109 Reference 109

Philo, J. S. (1994). Measuring sedimentation, diffusion, and molecular weights of small molecules by direct fitting of sedimentation velocity concentration profiles. In: *Modern analytical ultracentrifugation*. T.M.Schuster and T.M.Laue, eds. Birkhauser, Boston, pp. 156-170.

6.110 Reference 110

Schuck, P. (1998). Sedimentation analysis of noninteracting and self-associating solutes using numerical solutions to the Lamm equation. *Biophys. J.* 75, 1503-1512.

6.111 Reference 111

Moser, M. R. and Baker, C. A. (2021). Taylor dispersion analysis in fused silica capillaries: a tutorial review. *Anal. Methods* 13, 2357-2373.

6.112 Reference 112

Stetefeld, J., McKenna, S. A., and Patel, T. R. (2016). Dynamic light scattering: a practical guide and applications in biomedical sciences. *Biophys. Rev.* 8, 409-427.

6.113 Reference 113

Crocker, J. C. and Grier, D. G. (1996). Methods of digital video microscopy for colloidal studies. *J. Colloid Interface Sci.* 179, 298-310.

7 Index