INSTITUTO SUPERIOR DE ENGENHARIA DO PORTO

MESTRADO EM ENGENHARIA COMPUTAÇÃO E INSTRUMENTAÇÃO MÉDICA




Desenvolvimento de aplicação informática para a determinação de propriedades de difusão de químicos em tecidos biológicos

PEDRO MIGUEL FERREIRA PEIXOTO

Julho de 2015

POLITÉCNICO
DO PORTO

Instituto Superior de Engenharia do Porto

Instituto Politécnico do Porto

Desenvolvimento de aplicação informática para a determinação de propriedades de difusão de químicos em tecidos biológicos

Development of computer application for determination of chemical diffusion properties in biological tissues

Pedro Miguel Ferreira Peixoto

"Dissertation carried out under the scope of Master in Computer Engineering and Medical Instrumentation"

Adviser: Luís Oliveira (PhD)

2015

Pedro Peixoto, 2015

Acknowledgements

The writing of this thesis was possible thanks to the Instituto Superior de Engenharia do Porto.

Thanks to my adviser Luís Oliveira, for the patience, trust and guidance.

Last but not least, thank you to my family and friends that help me achieve my dreams and goals.

Resume

The determination of diffusion characteristics of chemical agents and drugs in biological tissue has become very important in the pharmaceutical industry and clinical investigation. As an example, the estimation of the diffusion properties of optical clearing agents in biological tissues allows the characterization of the optical clearing mechanisms involved such as tissue dehydration and refractive index measurement.

In clinical practice, such emerging techniques are also very promising. The study and discrimination of the diffusion coefficient between healthy and pathological tissues may allow to establish a diagnosis procedure.

The research in clinical related applications of optical technologies has grown significantly in the last decades.

Within biomedical optics, the study of chemical diffusion in tissues has recently showed great potential becoming evident the need for computer methods and applications to perform data processing and elaborated calculations, as described in literature.

Abstract

The main objective of this work was to develop an application capable of determining the diffusion times and diffusion coefficients of optical clearing agents and water inside a known type of muscle. Different types of chemical agents can also be used with the method implemented, such as medications or metabolic products. Since the diffusion times can be calculated, it is possible to describe the dehydration mechanism that occurs in the muscle. The calculation of the diffusion time of an optical clearing agent allows to characterize the refractive index matching mechanism of optical clearing. By using both the diffusion times and diffusion of water and clearing agents not only the optical clearing mechanisms are characterized, but also information about optical clearing effect duration and magnitude is obtained. Such information is crucial to plan a clinical intervention in cooperation with optical clearing.

The experimental method and equations implemented in the developed application are described in through out this document, demonstrating its effectiveness. The application was developed in MATLAB code, but the method was personalized so it better fits the application needs. This process significantly improved the processing efficiency, reduced the time to obtain the results, multiple validations prevents common errors and some extra functionalities were added such as saving application progress or export information in different formats. Tests were made using glucose measurements in muscle. Some of the data, for testing purposes, was also intentionally changed in order to obtain different simulations and results from the application.

The entire project was validated by comparing the calculated results with the ones found in literature, which are also described in this document.

Motivation

Understanding the diffusion characteristics of optical clearing agents in different biological tissues is extremely important. By studying, characterizing and discriminating between the optical clearing mechanisms designated as tissue dehydration and refractive index matching, the magnitude and duration of the effect to reduce light scattering are obtained. This information allows for effect optimization and selection in clinical applications. The process to achieve this information is complicated due to the data processing required of numerous treatments. The calculations are also extensive as well as complex due to the high number of measurements done in each treatment. Although calculations are extensive, this method has many fields of application. It can be used in clinical investigation, pharmaceutical industry, dermatology, cosmetic, diagnosis and treatment techniques. Given the increasing amount of studies from different fields and the complexity involved in the procedures, the creation of the developed application presented in this document is justified.

Index

List of Figures

List of Tables

List of abbreviations and symbols (alphabet order)

Abreviations

Adj R-sq - Adjusted R square

Coeff - Coefficients (regression coefficients, including the intercept)

Conc. - Concentrations

DFE - Degree of Freedom in the Error

OCA - Optical Clearing Agent

OCAs - Optical Clearing Agents

RI - Refractive Index

RMSE - Root Mean Squared Error

SSE - Sum of Squares due to Error of the fit

S.D. - Standard Deviation

UI - User Interface

Symbols

T_c Collimated Transmittance

λ Wavelength

1 Chapter 1

1.1 Introduction

In recent years a significant volume of research has been made to estimate the diffusion properties of optical clearing agents (OCAs) in biological tissues [1-2]. A simple and effective method to estimate the diffusion properties of OCAs in biological tissues based on the measurements of collimated transmittance (T_c) has recently been reported by different research groups [1-4]. The estimation of such properties provides the means for characterizing the optical clearing mechanisms involved in the clearing process [1-2]. Some study's even question how should these properties be used *in vivo* tissues [5]. There are also several soft and hard tissue in human and animal bodies as well as a great variety OCAs are known today. Researchers have used OCAs that are usually classified as: alcohols (polyethylene glycol (PEG) [6-7], butanediol [7-8], glycerol [8-10]); sugar alcohols (ethylene-glycol [1, 11], xylitol, sorbitol [7, 11]); sugars (glucose [2-3, 7, 12], dextrose, fructose, sucrose [7, 13-14]); organic acid (oleic acid [7]); and other organic compound (dymethyl sulfoxide [14-15] and propanediols 1.2 and 1.3 [11, 16]). These studies with different OCAs have the objective to create temporary , harmless and reversible transparency effects in different tissues. To study the diffusion time and diffusion coefficient for each particular OCA per tissue treatment is a way to describe the duration, and magnitude of temporary transparency created through the individualization of tissue dehydration and refractive index (RI) matching that clear the tissues.

This growth increases a demand of new tools to validate the results and improve the work of researchers involved, as well as increase their level of productivity. The hardware necessary to achieve this kind of work is usually at the professionals disposal, however, efficient software can be expensive or even non-existent, forcing the researchers to sometimes do elaborate

calculations by hand or design inflexible sequential processes that usually tend to be remodelled in every use or iteration, rather than make use of the capabilities of recursive methods, which are more efficient. This gap can be a step back on the work being done, because it consumes more time to the team involved in the research, since it must adapt the algorithm at every iteration.

In the course of this document it will be shown to you how the development and posterior use of agile software increases the work efficiency of certain researches. In this particular case, the calculation of diffusion characteristics (diffusion time and diffusion coefficient) with the experimental data collected by a spectrometer. The data used is the same as the one in the literature, so it could be compared to validate the application developed [2]. The application was developed in MATLAB due to the high quantity of tools and functions already implemented in MATLAB internal coding, that could and were successfully used in the developed application. In chapter 2 some of the tools and functions applied are described.

1.2 State of art in biomedical optics and diffusion characteristics

Conclusive studies are presented in literature relative to optical clearing effects using OCAs such as ethylene-glycol [1] and glucose [2]. One of the objectives is to better understand and characterize the time dependence of the optical clearing mechanisms. The first of these mechanisms is tissue dehydration (associated with water flux going out of the tissue) and the second is the RI matching (associated with OCA flux going in). This method tries to reduce the high scattering of turbid biological tissues that limits the penetration of visible and near-infrared light, in order to increase the image resolution as well as increase the light propagation through the tissue with minimal power loss [1-4].

There are a few reported diffusion times for some optical clearing agents such as fructose [17], glucose [2, 18-20], mannitol [19], glycerol [21], lactose [17] and dimethyl sulfoxide [22].

There are also reports on the use of new developed clearing agents like SeeDB (fructose, water and thioglycerol) used to clear the brain of a mouse, and capture images with a multi-photon microscope, which allows 3D visualization and reconstruction with the proper software. This agent clears grey and white matter with minimal deformation artefacts, alterations of volume or morphology (such as dendritic spines of pyramidal neurons in the cerebral cortex) of the brain tissue. The agent creates great transparency and the clearing time is also shorter than other agents [23-24].

Other optical imaging techniques were also developed to achieve deeper tissue imaging with second-harmonic radiation probes in living tissue, where 100 nm and 300 nm $BaTitO_3$ nanoparticles can be detected through the interval of 50 μm and 100 μm of mouse tail tissue (in vitro or in vivo) [25].

1.3 Theoretical methods

The computational work of the data collected is usually done through separated algorithms or methods rather than consolidated applications or software (with the exception of the software provided by the acquisition equipments such as spectrometers for example). Through out the literature some optical clearing methods require a great deal of computational power, as well as specific algorithms designed to manipulate the collected data, increasing the demand for certain computational and instrumentation breakthroughs so that it can be implemented in this field [26-27].

As described in these publications [26-27], a method based on Fick's law of diffusion can be used to study and discriminate the two individual fluxes relative to optical clearing mechanisms: tissue dehydration and RI matching. Tissue dehydration is the mechanism that occurs at the beginning of the treatment and takes the shortest time. It is associated with a unique water

flux out of the tissue, which is stimulated by the osmotic pressure created by the OCA in the immersing solution. Some diffusion time values for glucose and ethylene glycol were reported as less than one min [1-2]. The RI matching mechanism consists on the placement of the OCA in the interstitial location of the tissue and raising the RI of tissue fluids to higher values than observed in nature, better matched to the RI of scatterers [1, 26, 28]. The flux associated with this mechanism is the OCA flux from the immersing solution into the tissue, which takes longer time to occur. Some cases have been reported with with diffusion time values of about 5 min and 7-8 min for glucose and ethylene glycol diffusion in muscle, respectively [1]. This method involves elaborate calculations that use thickness and T_c measurements made during optical clearing to determine the diffusion time values of OCA in tissues [26-27].

This method allows the possibility of creating a T_c time dependence for various wavelengths within the spectral band limits of the spectra measured during treatments and these time dependencies can be fitted with lines that have exponential increase behaviour. Depending on the OCA concentration used in the immersing solution, the two fluxes that translate the two optical clearing mechanisms occur simultaneously, at least during initial time of treatment. The method described in literature can be applied, but the results of calculations for a mixed global flux does not allow mechanism discrimination. The inability of discriminating the two mechanisms of optical clearing is imposed by the fact that a particular and generic optical clearing treatments is associated with a mixed global flux of OCA going in and water going out of the tissue. To obtain the discrimination of the two fluxes, and consequently the discrimination of the two mechanisms of optical clearing, several treatments must be studied. A collection of treatments with different OCA concentrations in solution will allow establishing a relation between diffusion time of the mixed fluxes and OCA concentration in the immersing solution. The maximal value of this dependence translates the true OCA diffusion time in the tissue and the minimal diffusion time translates the diffusion time of water [1].

Equations involved in the calculations and method used to obtain the diffusion properties will be described in chapter 2.

2 Chapter 2

2.1 Development and implementation

Through out this chapter it is described the entire development of the application and the reasons for the implementation of certain methodologies to achieve the full length of the desired sequence described in the previous chapter. For this reason the following images and tables presented in this chapter do not reflect an experiment where the results of each step are connected. Instead, a series of tests performed during the implementation are presented, showing the relevance of the application limits and user choices that can sometimes induce cascaded errors. Some of the data, for this chapter only, has also been intentionally altered and modified in order to force certain responses from the application side. In the next chapter, a more detailed discussion of the results and expected values for an experimental study with glucose, is demonstrated.

2.2 Methods and algorithms implemented

The experimental method used to obtain the diffusion properties is based on T_c measurements and allows us to study and distinguish between the two diffusion fluxes, one for OCA and the other for water, that occur during the optical clearing of the tissue. To do this, the characteristic diffusion time and diffusion coefficient for OCA and water need to be estimated using Fick's Law of diffusion [1-3, 26-27, 29-30].

The process consists on immersing a slab-form tissue in aqueous solution containing a known concentration of a particular OCA. This arrangement allows the agent to diffuse into the tissue through both slab surfaces at the same time. This happens due to the difference of agent concentration between the tissue and immersion solution and also because the agent has

osmotic diffusion properties. Such diffusion can be explained mathematically with Fick's Law [26-27].

$$\frac{dC}{dt} = D\frac{d^2C}{dx^2} \qquad (1)$$

Equation 1 characterizes the time dependence of the agent's concentration C in any possible position x (perpendicular direction to slab surfaces) inside the tissue thickness. The diffusion coefficient for the flux being studied in equation 1 is D but it is sometimes presented in literature as D_a [26], because the diffusion coefficient is related to an agent (same thing happens in the notation used for concentration C and C_a when related to an agent).

The diffusion coefficient and the diffusion time of the agent are related through equation 2, where τ represents the diffusion time, and d is the thickness of the tissue sample at that given time.

$$\tau = \frac{d^2}{\pi^2 D} \qquad (2)$$

The immersing solution should have a higher volume than the tissue to guarantee that there is always an excess of OCA in the solution and a free, or almost free, diffusion to the agent into the tissue. The amount of dissolved matter m_t in the tissue at a given time t relative to its equilibrium value m_∞ can be determined by equation 3. A ratio of the volume averaged concentration at a particular time, inside the tissue, is presented in equation 3.

$$\frac{m_t}{m_\infty}=\frac{\int C(x,t)dx}{C_0}=1-\frac{8}{\pi^2}[\exp(\frac{-t}{\tau})+\frac{1}{9}\exp(\frac{-9t}{\tau})+\frac{1}{25}\exp(\frac{-25t}{t})+...] \quad (3)$$

The relation between the time dependence of the OCA concentration within the tissue and the characteristic diffusion time, which is a first order approximated solution to equations 1 and 3 is presented on equation 4.

$$C(t)=\frac{1}{d}\int C(x,t)dx\simeq C_0[1-\exp(\frac{-t}{\tau})] \quad (4)$$

To estimate the diffusion properties of an OCA in a biological tissue with these equations it is necessary to use T_c measurements made from a sample under treatment with an aqueous solution of the agent.

These tissue measurements are made using the assembly presented in figure 2.1, where the sample is fixed in a horizontal position and the transmitted beam passes through the sample. In effect, to allow flux discrimination between optical clearing mechanisms, several treatments with different OCA concentrations in the immersing solution are necessary. The experimental methodology described bellow is applied to all treatments. The structure that maintains the sample fixed has a circular form, with a central 1mm hole so that light can pass through. The immersing solution is injected after the natural T_c spectrum is measured from the sample. Treatment measurements are initiated as the sample is immersed in the solution and natural T_c is considered to be measured at t=0s.

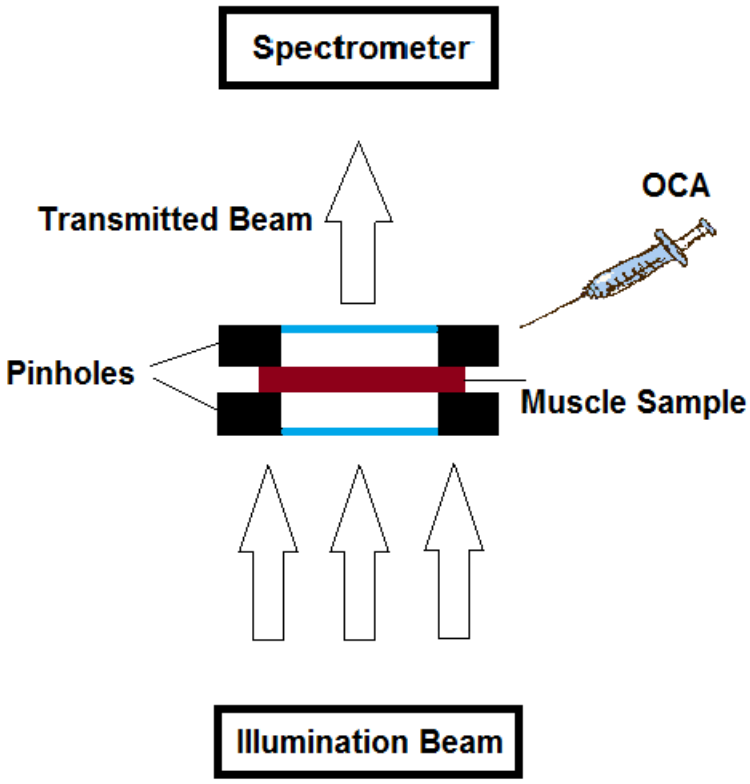

Fig. 2.1 Usual assembly to measure T_c *values.*

This measuring process is repeated as many times as the number of experiments desired. In each particular treatment a new muscle sample is to be used under treatment with a new OCA concentration in the immersing solution. To avoid treatment contamination, the entire assembly must be cleaned between experiments.

The measurements are then collected and processed. The time variation of T_c for different wavelengths is then represented for every treatment. Those selected wavelengths are usually selected within the bandwidth where light scattering dominates over absorption for that type of tissue. In the case of skeletal muscle used in measurements considered in this work, such scattering dominating band is between 600 and 800 nm [1].

Light scattering occurs due to the existence of irregularities in the light path, such as particles and different interfaces with unmatched RI. The scattering phenomenon is usually categorized in two large classes, strongly scattering

(opaque) and weakly scattering (transparent) [31]. By selecting such bandwidth, and contaminating the interstitial space of the tissue with an OCA, an attempt to match the RI of the tissue components in order to obtain better results (deeper light propagation) is made.

With the ideal bandwidth selected to implement this study, the entire T_c time dependence for each λ is displaced vertically so the natural T_c is set to 0. Then, the beginning of the saturation regime (regime where no significant flux is observed) is identified and the data for longer time than this is neglected. The displaced dataset delimited by the beginning of the saturation regime is normalized to the highest value so the entire dataset ranges from 0 to 1.

In order to obtain the diffusion time of the combined fluxes in each treatment the normalized and displaced data are fitted with a curve with the form of equation 5.

$$T_c(\lambda) = 1 - \exp\left(\frac{t}{\tau}\right) \qquad (5)$$

Equation 5 results directly from equation 4 and it is a sensitive method to evaluate OCA concentration within the tissue from T_c measurements. When fitting each dataset correspondent to a particular wavelength with a line described by equation 5, a diffusion time value is obtained. For a particular treatment several time values are obtained, one for each λ within the considered band [1]. Considering the various treatments studied, the final step of the estimation methodology consists on representing the mean diffusion time values as a function of the OCA concentration in the immersing solution. Such representation provides valuable information for the researchers. First of all, if the OCA concentrations were correctly selected, it is expected to obtain a peak in this representation. This peak corresponds to optimal OCA diffusion into the tissue due to the equilibrium verified between the water in

the immersing solution and the free water in the tissue [1]. By identifying this peak its possible to obtain the free water content on the tissue from the x-axis of the graph and the OCA diffusion time from the y-axis. On the other hand, from the treatment with a highly saturated immersion solution (the highest saturated as possible), the water diffusion time is obtained from the y-axis of the graph. The diffusion time values of OCA and water allow discriminating the two optical clearing mechanisms and can be used in equation 6 to calculate the correspondent diffusion coefficients for these fluids.

$$D = \frac{d^2}{\pi^2 \tau} \qquad (6)$$

With equation 6 the diffusion coefficients (D) can be determined both for water and OCA by using the respective diffusion time values (τ) and correspondent sample thickness (d) value.

2.3 Reading initial data and consequent organization

The developed application only reads files in *txt* format, with the λ and T_c separated by a tab and generated automatically using the spectrometer software (AvaSoft from Avantes™ for the measurements used in this particular work). The initial data provided consists in individual files, separated by OCA and concentration folders with T_c measured per λ at a given time, and given OCA concentration in the solution used. This data needs to be correctly loaded into the application so that the methodology previously explained can be applied without errors. Figure 2.2 shows data from the experimental measurements made during treatments with glucose solutions that were used in this work. Treatments with eight concentrations between 20% and 54% were made and measurements of T_c and samples thickness were performed during

each of this treatments. Note that the OCA name must be in the file name as well as the respective concentration. As an example, Glucose 250 corresponds to a concentration of 25%-glucose in the solution used to treat the tissue.

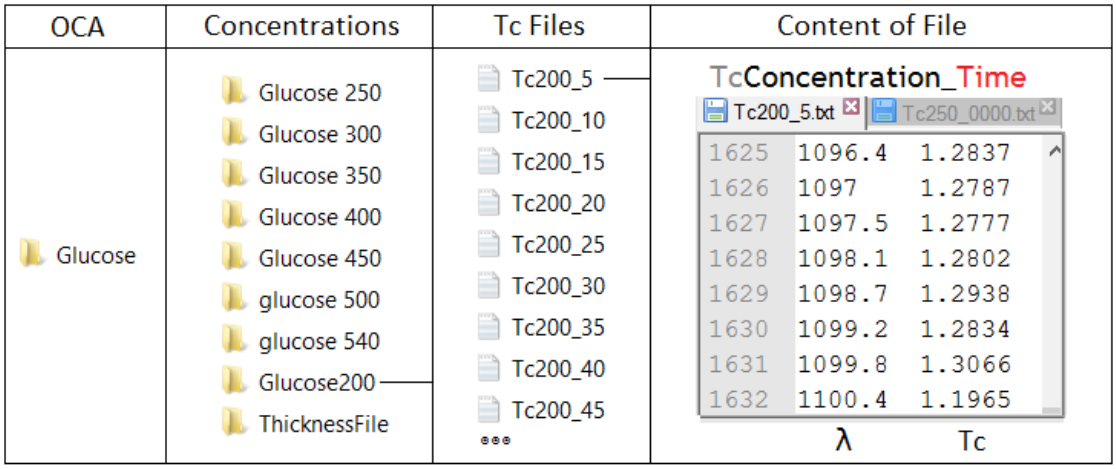

Fig. 2.2 Typical organization of the initial data.

By taking a closer attention to figure 2.2, the initial data may sometimes have inconsistent file naming formats as well as corrupt data. However, the application can surpass some of those misleading names.

Inside the concentration folders, each file contains two separated columns (separated by a tab). The λ values which are common in all the individual files over time, and the T_c values measured by the spectrometer. Naming each file accordingly reveals itself useful (ex: Tc200_n.txt is renamed to Tc200_0000.txt) for the remaining steps of the application. The first letters represent the type of measurement, T_c values, followed by numbers representing the concentration of OCA used in the solution to treat the tissue, in this case it is 20%. Finally, at the end of the file name and separated by an underscore there is a time instant of treatment (in seconds) that the measuring of all the values, from the current file, occurred. The initial data, once again, had different formats to indicate time, using the letter *n* for natural tissue which represents the instance of time 0s, as well as only one zero or four zeros (ex: Tc200_n.txt, Tc200_0.txt or Tc200_0000.txt), where any of the cases correspond to the same instant of time.

In order to surpass this disagreements all files are checked and saved in a temporary folder so that all the files loaded to memory have a consistent format and are read correctly in the order of time they were measured.

This change consists in modifying all file names to the same length, so that they are read accordingly, in order of time, instead of shorter length name first. The letter *n* (natural tissue) is replaced by four numbers representing the time of treatment at 0 s (see top of third column in figure 2.3). The reason for using four numbers to describe the time of treatment is related to the need of representing the maximum value, which is 1800 s (for 30 min treatments, as considered in experimental methods used in this work). As a precaution measure the original files remain intact, and only the new files saved in the temporary folder suffer any kind of change during the use of the application.

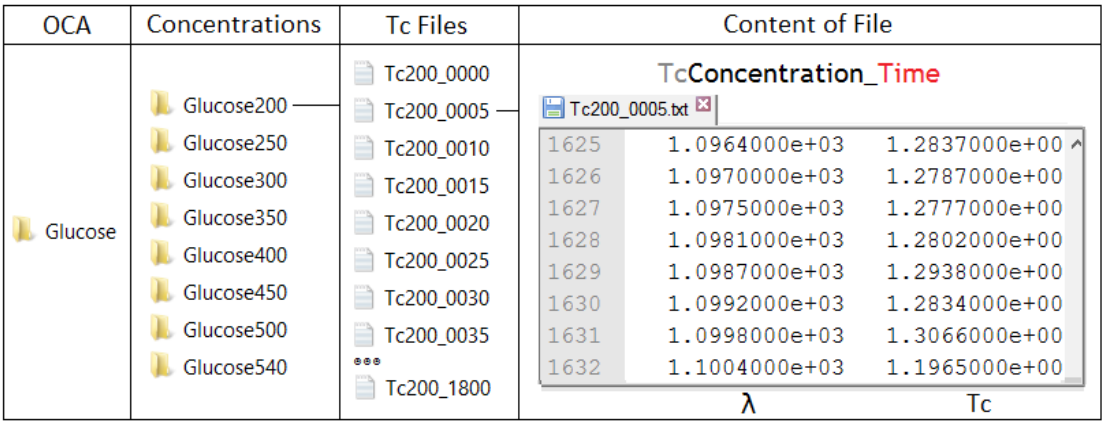

Fig. 2.3 Organization of an OCA folder in the temporary folder.

Access to the functions that can read, write, delete or change in the temporary folder with the use of the developed application requires a series of validations in order to prevent the malfunction of the methodology described.

The temporary files also work as a safeguard if the application stops unexpectedly, so that they can be used again by the application. But the main reason for this organized content is that the developed application can create

a matrix with all the T_c values, a vector with all the λ values, and a vector with all the times of treatment obtained through the name of each file. These are three main variables that the application systematically uses, and two of them are common in all the concentrations, the vector that contains the λ that corresponds to the first column of every content file from figure 2.3, and the vector with all time instants that corresponds to the last numeric part of each T_c file. By taking a close attention to the third column named Tc Files from figure 2.3 and only read the numeric part to the right of the underscore, a vector with all the time instants can be formed.

Table 2.1 has T_c values as a function of λ and time.

Table 2.1 T_c measurements with 20%-glucose solution.

		Wavelength (nm)								
		171.49	172.06	172.63	173.2	173.77	174.34	174.9	...	1100.4
Time (s)	0	1.29	1.35	1.42	1.39	1.42	1.38	1.35	...	0.84
	5	1.46	1.58	1.56	1.54	1.53	1.52	1.52	...	1.2
	10	1.67	1.95	1.93	1.9	1.91	1.86	1.86	...	1.42
	15	1.55	1.79	1.9	1.9	1.91	1.86	1.84	...	1.59
	20	2.36	2.4	2.31	2.27	2.27	2.22	2.19	...	1.82
	25	3.15	2.94	2.95	2.82	2.82	2.8	2.76	...	2.17
	...	...	...	...	...	...	...	...	...	...
	1800	2.9	2.98	3.13	3.09	3.09	3.06	2.99	...	2.96

The first column on the left side represents the sequence of time instants and the first row the λ values. The complete content of this table can be exported directly from the developed application into *xlsx* or *txt* files. Besides

the generated table 2.1, the application can also show and export a graph for each concentration; a figure such as 2.4. This allows the user to explore the initial data by evaluating the variations of the T_c over time and λ not only by consulting a table but also through the visualization of a graph. Each continuous line in figure 2.4 is the measurement at one given time of treatment.

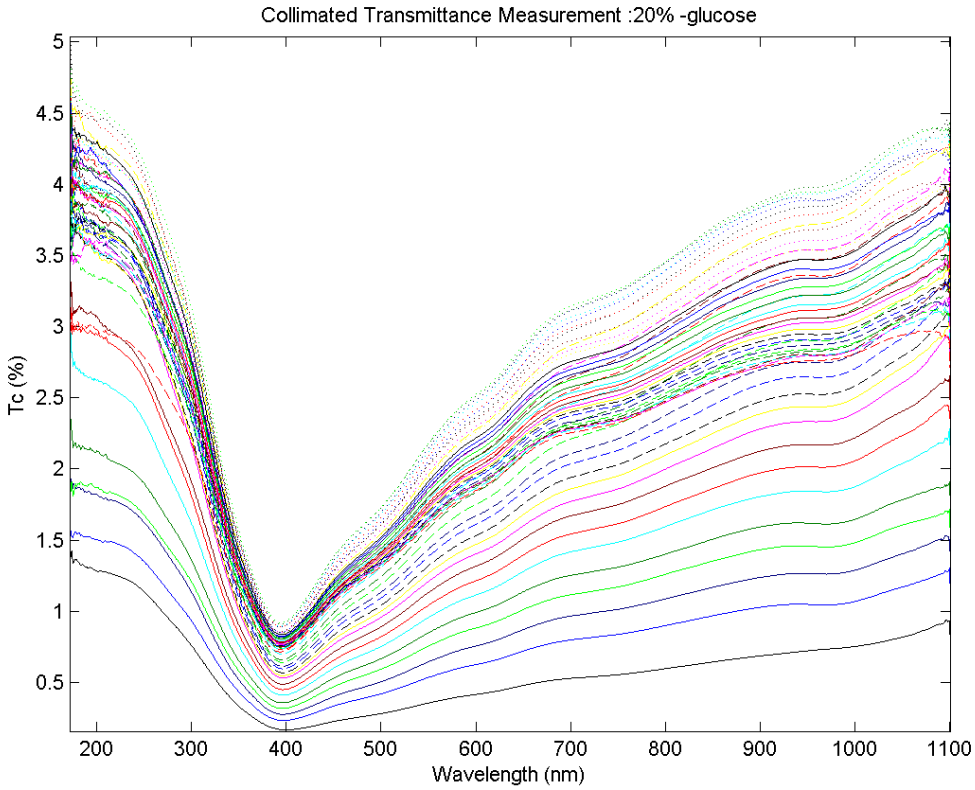

Fig. 2.4 T_c measurements with 20%-glucose.

Although the bandwidth of interest for the calculations to be made by the application is between 600 and 800 nm, figure 2.4 presents the entire acquisition bandwidth of the spectrometer. This is raw measured data before any processing.

2.4 Bandwidth selection

The λ range of the illumination beam can be widely spread, as well as the acquisition band of the spectrometer used to measure the T_c spectra. However, the bandwidth of interest is usually smaller and its range can be reduced in order to minimize the resources needed to manipulate the data. According to literature, for skeletal muscle, the scattering coefficient is significantly higher than the absorption coefficient between 600 and 800 nm [5, 26, 28, 32-34]. This means that the selection of a specific bandwidth is clearly important and expected by the user's point of view and this functionality is achieved by the correct manipulation of the data saved in the temporary folder, through the application interface. Furthermore, for future use of this application in other studies, other biological tissues with different optical properties and different λ band with dominating scattering can be considered.

Due to the bandwidth selection, the application has a series of validations to prevent the user from inserting incorrect inputs. As well as a method to calculate the closest values to the ones desired by the input interval and respective spacing between each λ value. As an example, in table 2.2 the bandwidth range is between 400 and 1000 nm spaced by 100 nm. This means 7 wavelengths.

Table 2.2 The desired bandwidth is the first line (Input), the second line (Output) is the closest value of each wavelength.

Input	*400*	*500*	*600*	*700*	*800*	*900*	*1000*
Output	*399.87*	*500.1*	*599.77*	*700*	*800.24*	*899.9*	*1000.1*

The bandwidth intervals can have floating numbers as inputs, but the number of wavelengths can't be a floating number. Figure 2.5 shows the application log giving a warning that those parameters are invalid, and cannot proceed.

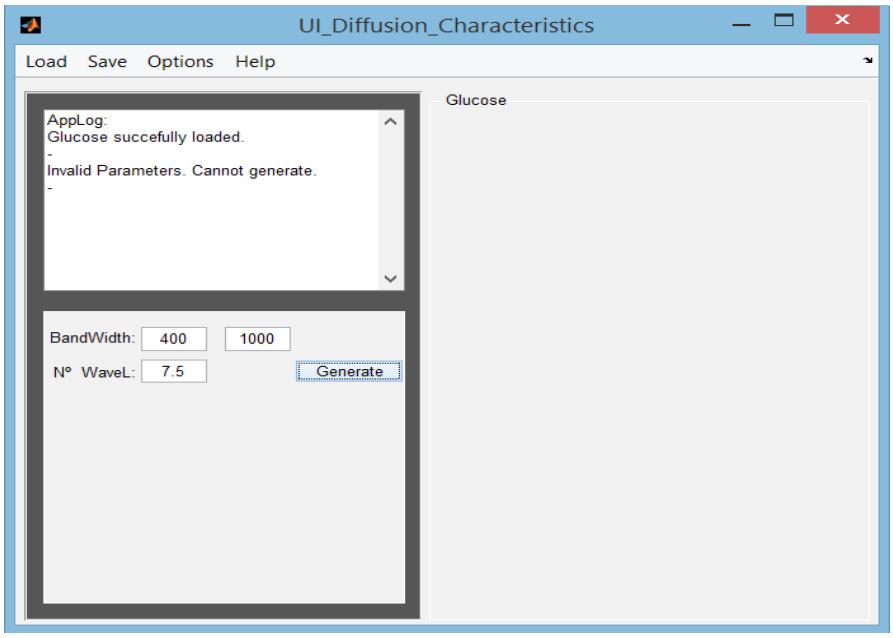

Fig. 2.5 Developed application interface. Bandwidth selection is available after an OCA is correctly loaded.

However, some Input λ values may not exist in the common vector of wavelengths (the common vector is the first column of the content file of figure 2.2 or 2.3), so the application returns the closest value for each λ value. For $\lambda = 800\,nm$ the highest percentage error can be found in table 2.2.

$$Percent\ Error = \left| \frac{800.24 - 800}{800.24} \right| * 100 = 0.03\,\% \qquad (7)$$

By taking a closer look over the common vector containing the wavelengths, the spacing between each one is around 0.56 nm (spectrometer slit), this means that the maximum percentage error is 0.035%. Then the larger the spectrometer slit is, the greater the error. When the selection of the bandwidth and spectrometer slit is made unwisely (when the Input values are too far from the output). If the user wishes to minimize this error then the spectrometer slit must be adjusted so the error generated by the selection of the bandwidth is as low as possible.

2.5 Maximum detection

The T_c measurements to be used in calculations with the application correspond to various treatments with different OCA concentrations. Depending on OCA concentration in the immersion solution used to treat the tissue sample, the beginning of the saturation regime can occur sooner or later within the treatment time. Since the active water and OCA fluxes that provide the optical clearing mechanisms occur before the beginning of the saturation regime, it is important to identify the time of treatment when saturation begins. As an example figure 2.6 shows the treatment with 40%-glucose, where it can be seen a smooth increase in T_c during the entire 30 min of treatment. Such fact indicates that saturation is only seen at the end of 30 min treatment.

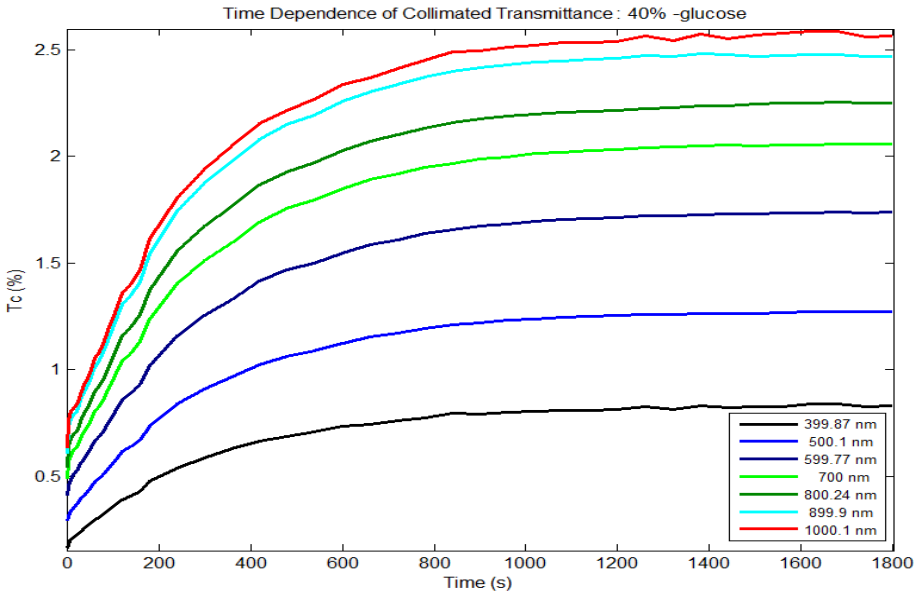

Fig. 2.6 T_c time dependence for muscle treated with 40%-glucose.

After each step of the application functionality, such as the initial data reading or the bandwidth selection, the temporary folder is refreshed, saving the relevant variables in the folder so that they can be used on the following functionality.

Using the selected bandwidth data saved in temporary files, the detection of the maximum of each curve that corresponds to an individual λ from the concentrations can be done automatically. However, this automatic detection should only be used as a starting guideline to find a common maximum for all curves in a given concentration of OCA, and there are a few reasons for that. The concept of detecting the highest number from an array is easy to implement when that maximum is unique in that array, but when several positions (different indexes) in the same array have the same maximum the returned index indicating the maximum value will vary, depending on the method used. The application uses only one method to find the maximum in an array (each curve that corresponds to a particular λ) and consists in returning the position of the first maximum value found (considering that the array has different positions with the same maximum value). Figure 2.7 shows a sine wave where multiple maxima exist in the same function, another similar situation is a straight line between two points where all the points of that line are maxima.

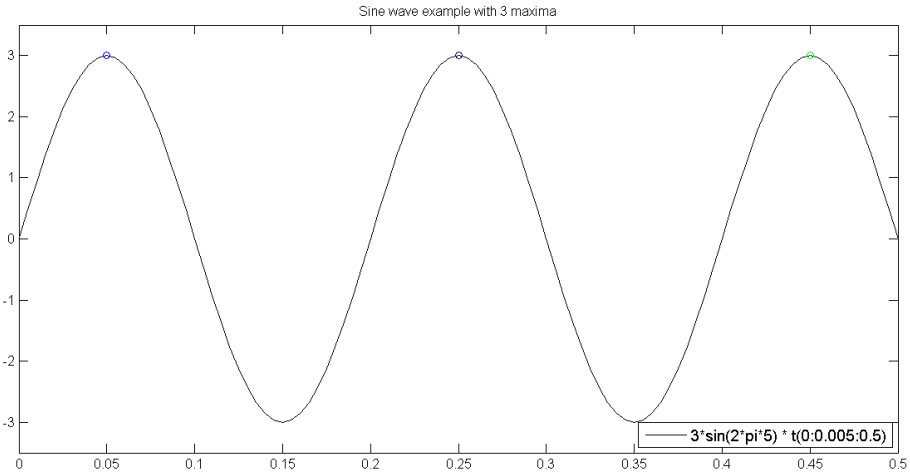

Fig. 2.7 Sine Wave example demonstrating multiple maxima in the same function.

Another reason to consider the automatic selection as a guide line is because inside the same concentration the returned maxima of each λ are usually different, as can be seen in the figure 2.8, and it is necessary to select a common maximum.

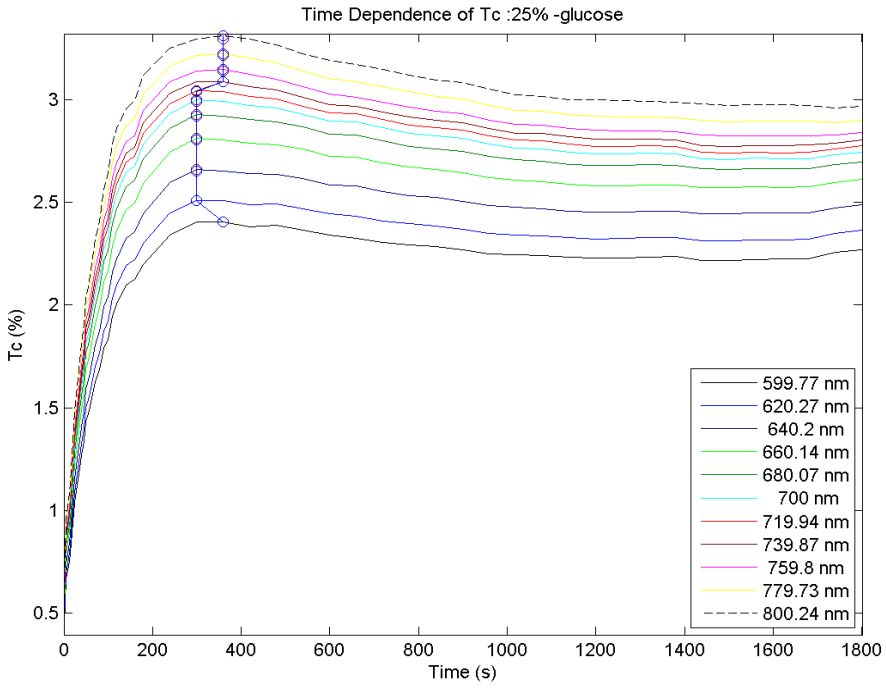

Fig. 2.8 Time dependence of muscle T_c treated with 25%-glucose, bandwidth[600-800], 11 wavelengths. Each λ has one maximum.

Once again the results vary depending on the method used, and it is not clear what method is the most adequate to apply because not all concentrations have the same form. The second part of the method used by the application is the mean of all maxima (obtained from all curves), followed by an approximation to the nearest value, since the result may not exist (for example the result gives 285.5 seconds, but it doesn't exist, so it rounds it up to 300 seconds). In figure 2.8 all the maxima are relatively close to each other (in order of time, x-axis) as most of the concentrations tested in the application but there are some concentrations where the difference, in time, is over 1000 seconds, which is more than 55% of the total time. Different methods could be applied here to find the individual maximum of each line such as finding all peaks, points in function where the derivatives are 0, first or last maximum, followed by a second part to find a single common maximum to the entire concentration which would involve the most common averages such as mean, median or mode. However, it would not remove the

need from the user of verifying the results, and most likely change them manually.

Using the automatic maxima detection as a guideline, the user can change maxima values through a editable table made visible after the button is pressed like the one shown in figure 2.9.

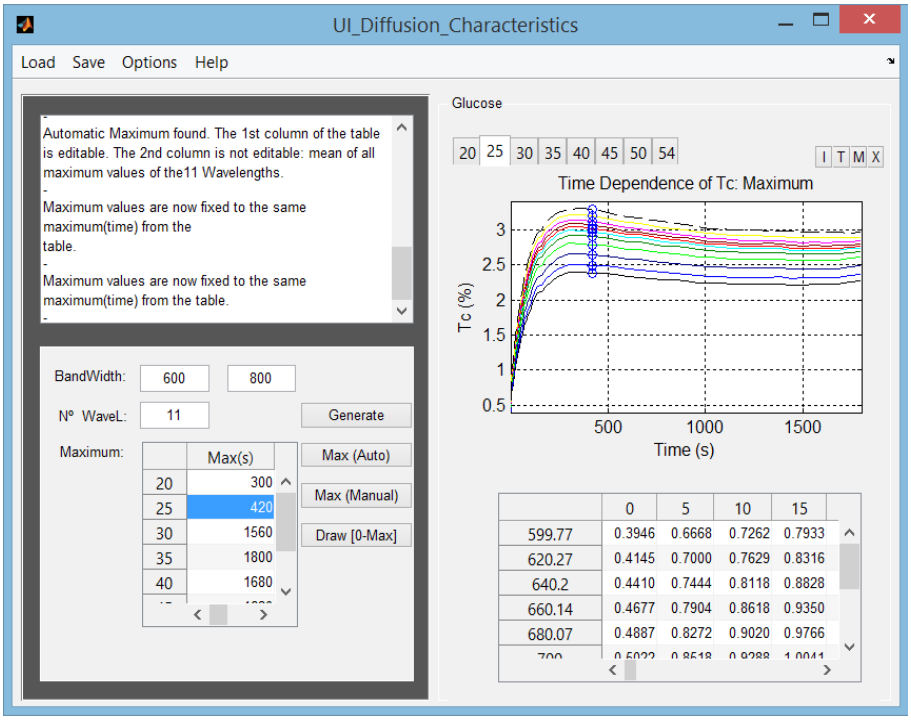

Fig. 2.9 Application Interface. The table next to the Maximum label is editable, allowing the user to change the maxima values of all concentrations.

By editing this maximal values of each concentration the previous maximal values are overwritten when the respective button (manual maximum) is pressed, as well as necessary validations take place in order to prevent incorrect data to be inserted, such as letters, symbols or none existent time values. If the user tries to change the number of wavelengths or bandwidth the application will make all the maxima buttons and the editable table invisible again, forcing the user to repeat all previous steps, preventing any kind of errors.

The next step after confirming the new maximal values is to consider only the data from the beginning of treatment until the time observed for the maxima. Once again the relevant data is always saved in the temporary folder.

2.6 Curve fitting and respective adjustments

A preparation process of the re-organized data in the previous step is necessary so that it is possible to fit the time dependencies with curves using equation 5 found in literature [1-2, 26]. These adjustments consist in shifting all the T_c values on the y-axis, so that the first T_c measurement is 0, when the tissue is in its natural state (t=0 s), and normalize the shifted data, as previously described.

Figure 2.10 represents a maximized window that is presented when the M button from the application (next to the X button that closes all tabs, at north-east) interface is pressed.

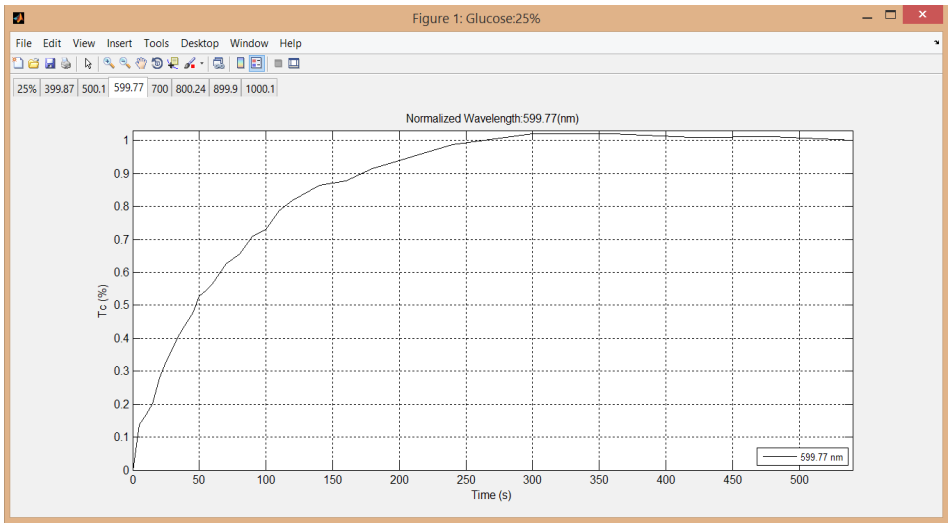

Fig. 2.10 Maximized window of shifted and normalized time dependence of T_c values.

Figure 2.10 presents shifted and normalized dependence of T_c at 25%-glucose with a bandwidth ranging from 400 to 1000 nm and 7 wavelengths. The selected tab displays the T_c time dependence for $\lambda = 599.77\,nm$.

Each tab in such window represents the various curves corresponding to all wavelengths for a particular treatment. Each of these curves can be seen individually by selecting the corresponding tab. Figure 2.10 is an example for $\lambda = 599.77\,nm$ (name of tab corresponds to the λ curve). Other tabs have similar representations for other wavelengths. A tab with the entire collection of curves corresponding to all wavelengths for a particular treatment (25%-glucose in this case, is the first tab on the left) is also present in the maximized window.

The data presented in figure 2.10 has a false maximum. From the graph its possible to see that after displacing and normalizing the experimental T_c data, some of the values in the curve are higher than unity. This is observed for data on the curve after 300 s, which can also be consulted through table 2.3.

Table 2.3 Shifted and Normalized T_c values, at 25% glucose, bandwidth[400-1000], 7 wavelengths.

		Time (s)						
		...	240	300	360	420	480	540
Wavelength (nm)	399.87	...	0.9640	1.0172	1.0138	1.0051	1.0213	1
	500.10	...	0.9828	1.0287	1.0244	1.0066	1.0130	1
	599.77	...	0.9800	1.0274	1.0204	1.0088	1.0114	1
	700.00	...	1.0014	1.0287	1.0260	1.0179	1.0132	1
	800.24	...	1.0096	1.0274	1.0329	1.0263	1.0162	1
	899.90	...	0.9977	1.0176	1.0305	1.0299	1.0109	1
	1000.10	...	0.9760	1.0129	1.0341	1.0330	1.0142	1

The reason for this type of error to occur is the bad choice of the maximum in the original measured data. Instead of identifying the maximum at 360s, it was wrongly selected at 540s. This example shows the importance of selecting and reviewing the maximum for each curve so that eventual errors can be avoided. In this example the user should consider changing the maximum value for this treatment, in order to minimize the generated error caused by the wrong index selection. However, in some cases it is impossible to prevent this error, as mentioned in the previous step.

Regarding experimental data fitting, MATLAB has a curve fitting tool (cftool) that allows adjusting experimental data with different curves described by various equations. For a particular treatment studied, by pressing the CF Tool button, the application opens several tabs in a new window. Each of these tabs in the new window represents the dataset for a particular λ in the time interval delimited by the occurrence of the maximum considered before. Such representation in each of the tabs is made using dots and shows the T_c data after displacement and normalization. The user will then select the appropriate fitting to create the curve. To do this, the user must select the "custom equation" option from the drop-down menu. By doing this and writing the custom equation according to equation 5, cftool will create and represent a fitting curve to the data-points. When the curve is created, the fitting statistics are presented in the window on the left (results), and bottom (table of fits). The user can change the method and algorithm that characterize the fitting curve through the fitting options. A combination of the various options will give different results.

This process is repeated for the other wavelengths, by creating the desired fitting curve in each tab from the cftool window.

Since the objective of the user is to obtain a mean diffusion time for each treatment with a particular OCA concentration, the fittings for various wavelengths must be performed for each treatment, one at a time.

Cftool is the only component of the application that wasn't developed by the author. It uses more resources than initially expected, because of the high variety of properties that can be adjusted and different functionalities contained inside this tool. But it does offer very satisfying results. Figure 2.11 demonstrates the fitting statistics, using equation 5, to fit the curve.

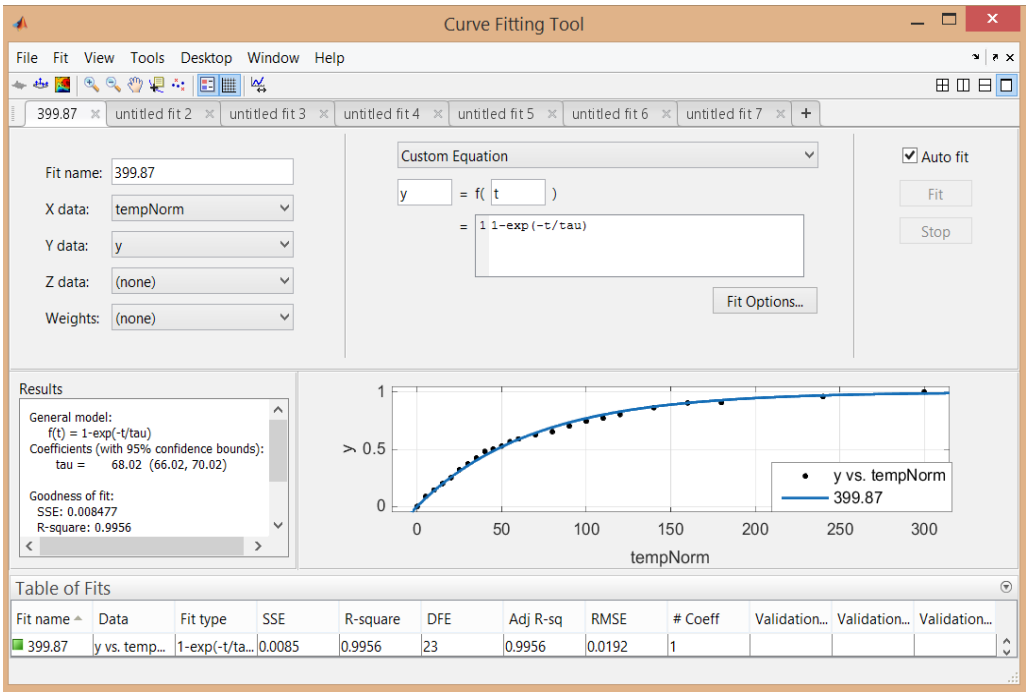

Fig. 2.11 Curve fitting tool window with various tabs containing the experimental datasets for each wavelength within a particular treatment.

The tabs should be named after the wavelengths generated in the bandwidth. The data is always loaded in ascending order of the wavelengths, and one concentration at a time.

When the curve fitting tool is opened by selecting the desired concentration through a list-box and pressing the respective button, an auxiliary window should also be opened to allow the user to continue the sequence of the described methodology in chapter 1. In this auxiliary window (represented in figure 2.12) there is an empty editable table, with the OCA concentrations and the wavelengths generated through the selection of the bandwidth, as

well as equation 5 which can be copied to the curve fitting tool as a custom equation.

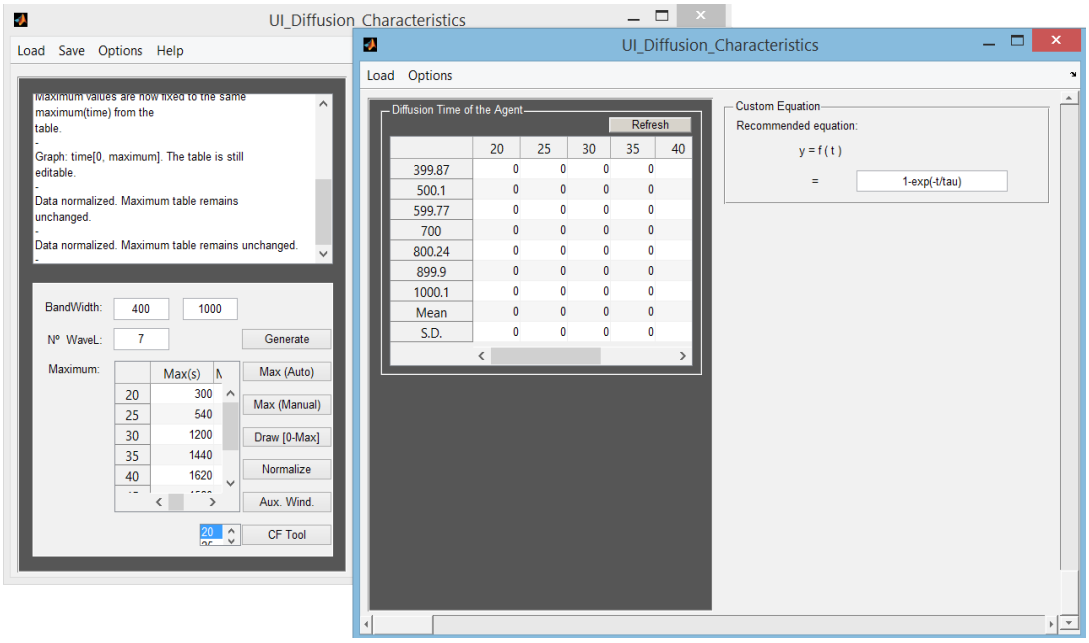

Fig. 2.12 Application windows. The auxiliary window is to the right.

The first row on the table contains all the OCA concentrations used in the study, the first column has the generated wavelengths, the mean value of the diffusion times and its standard deviation (S.D.). Each editable column of the table, from figure 2.12 represents a concentration of OCA, that must be filled with the calculated tau values found on figure 2.11 (inside the multi-line text-box named Results located south-west). Each tab from curve fitting tool corresponds to one λ (one editable cell under the respective concentration) from the table located in the auxiliary window.

When opening the auxiliary window, it loads an empty table with concentrations, wavelengths, respective mean and S.D. When opening cftool the respective concentration, or treatment, can be selected through the drop down next to cftool button in figure 2.12 (left window).

After the first column is filled, corresponding for example to the case of the treatment with 20%-glucose with 7 wavelengths between 400 nm and 1000 nm, refreshing the content will calculate the mean diffusion time of the

wavelengths, and the respective S.D., as well as save the content of the table into the temporary folder. In the next chapter it is described which parameters in the curve fitting tool (Figure 2.11) should be taken more seriously in order to manipulate the curve and alter the results.

This entire process of calculating the diffusion times is quite demanding. Since there are several resources in the fitting process that need to be adjusted to obtain an optimal fit, and due to the fact of several fittings are needed for each treatment, this whole calculation process is highly time consuming. The lack of experience required to fit the curves and obtain reasonable results through the fit options inside the tool can also increase the time necessary. For this reason it is strongly advised to save the curves relative to each treatment, not only to allow the resume of the work in another day, but also to keep the properties used inside the tool as reference points to replicate the methodology or experiments, since different options can give very different results (cftool allows the entire session to be saved, and resumed). The initial window (left window from figure 2.12) should remain open. This way it is possible to generate the next concentration with the curve fitting tool, after the respective column is filled with the tau values from the results of the curves.

2.7 Diffusion Time

One of the diffusion characteristics of OCAs in biological tissues presented in literature is the diffusion time, which is obtained for a particular concentration of OCA in the immersing solution. Please note that such particular concentration may not be one of the particular cases studied experimentally, but it can be estimated from calculations. The OCA diffusion time is obtained for a particular concentration that has the same water content as the free water in the biological tissue. This way, by determining that particular OCA concentration in solution, it was also estimated indirectly the amount of free water in the tissue, which is unknown many times [1-2]. The

final step of the implemented method consists on representing the calculated mean diffusion time as a function of OCA concentration in the immersing solution. This collection of graphical data points are fitted with a smooth spline function to obtain dependence between diffusion time and OCA concentration. By constructing such fit to the mean data, unknown diffusion time values for other OCA concentrations are interpolated. Such method detects the maximum OCA diffusion time, all peaks (if more than one exist), maximum concentration used and minimum diffusion time. All these values are important to detect eventual error and the diffusion time values that correspond to the optical clearing mechanisms: tissue dehydration and RI matching. Figure 2.13 shows the graph with mean diffusion time as a function of OCA concentration, the fitting spline and an erroneous secondary maximum (a second peak).

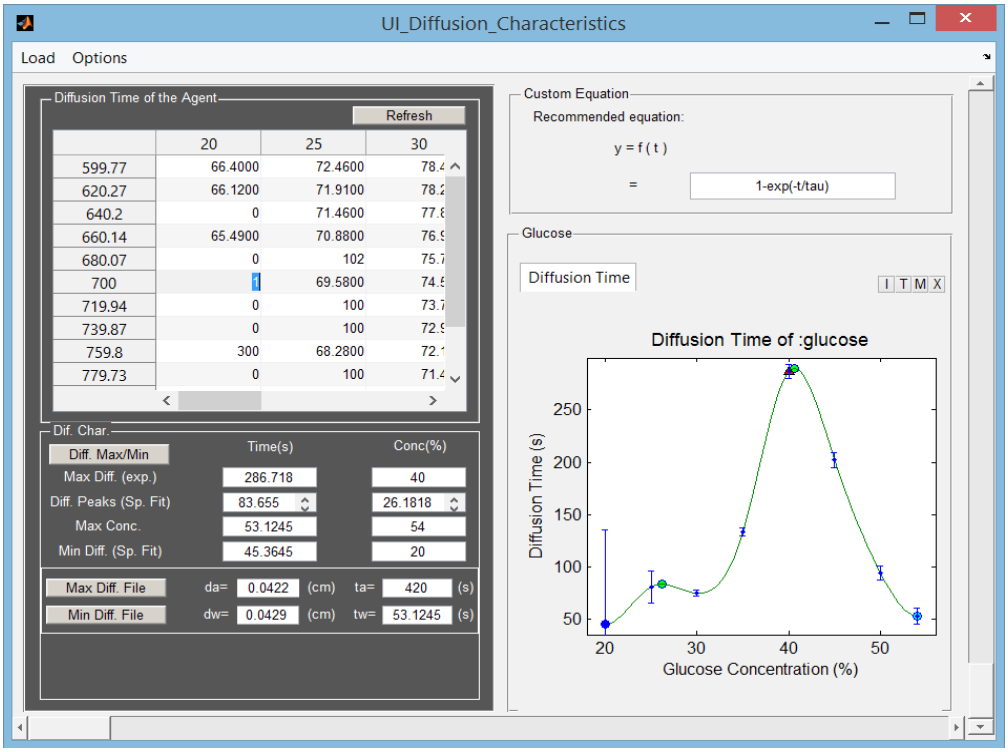

Fig. 2.13 Application Interface with diffusion times table filled and respective smoothed polynomial function, with 2 peaks.

The green dots presented in figure 2.13 represent the local maximum values found in the function. The reason for showing all peaks and not only the

absolute maximum is to point out additional inflexion points, slope inversions, that are uncommon in the usual graph format of diffusion time over concentration since, usually, the function should only have one peak (absolute maximum). The red triangle presented in figure 2.13 is the highest diffusion time that was obtained experimentally. It is not exactly the same as the spline maximum, but it should be very close, depending on OCA concentration selection for the study. The user must always compare these two values in order to evaluate if the treatments in use were selected with the correct concentration of OCA. Meaning that one of the treatments is expected to be the optimal OCA concentration in the immersing solution. In this particular case it is very close to 40%.

For the case of the muscle it was observed that 40.5% is the ideal concentration of OCA in the immersing solution [1-2]. Such value indicates that the free water content in the skeletal muscle is 59.5% (100% - 40.5%).

The difference between the experimental maximum (red triangle) and the spline maximum (highest green dot) can be significantly large in some experiences, especially if the difference between OCA concentrations used is too big. Leaving this choice to the user, will force him to evaluate the difference, and consider to perform new measurements if necessary.

The light-blue dot indicates the diffusion time that corresponds to the maximum OCA concentration used in the experiments. Since the maximum concentration is always the last experimental point in the function, it is easily identifiable. The dark-blue dot is the minimum diffusion time of all concentrations, and sometimes is not seen in the graph because it is overlapped with the time at maximum concentration, however in this case they were made distinguishable.

The reason why it is so important to obtain the diffusion time for the treatment with the highest OCA concentration is because the solution becomes saturated with OCA. Due to a much higher OCA concentration in the solution a very strong osmotic pressure is created over the tissue sample at the beginning of the treatment, forcing it to dehydrate fast, before OCA

diffusion into the tissue is significative. This means that the diffusion time obtained for a saturated solution corresponds only to the dehydration mechanism of optical clearing [1].

2.8 Diffusion coefficient

Once determined the optimal OCA diffusion time and water diffusion time, the calculation of the corresponding diffusion coefficients can be done when the required files containing the thickness values are available. If these files are not available the values can be introduced manually.

On south-west of figure 2.13 there are spaces to fill the thickness data of the tissue sample that correspond to the treatments with the optimal and highest concentrations of OCA. These values are used to calculate the diffusion coefficients of OCA and water, respectively. They do characterize the OCA and water fluxes that correspond to the tissue dehydration and RI matching mechanisms [1-2].

Table 2.4 represents the usual structure of the thickness files which can be loaded to calculate the diffusion coefficients.

Table 2.4 Typical data of a thickness file.

OCA Concentration-40%	
Time (s)	*Thickness (mm)*
0	*0.5000*
15	*0.4770*
30	*0.4623*
...	...
1800	*0.4350*

The labels can't be included in the file (Concentration, Time and Thickness) and the files must be saved as *txt*.

The content of the loadable file are the values in table 2.4 which were measured in mm, but according to equation 6 the thickness (d) is in cm, so the necessary conversion is done internally. If the loaded file has the tissue thickness values in another measurement unity (different from mm) it is necessary to compensate those changes, since the initial application is expecting the content of the data in mm to perform the consequent conversion to cm. Same thing goes for time, which must be in seconds.

The diffusion time of OCA inside a tissue sample is obtained by finding the optimal diffusion time (peak value retrieved from graph in figure 2.13), which as stated before is somehow near some experimental value.

In south-west corner of figure 2.13 there is a group with some input text boxes to introduce the optimal diffusion time of OCA (ta in figure 2.13) and sample thickness (da in figure 2.13) at that specific time and treatment. By manually introducing these values, the diffusion coefficient of OCA is calculated with equation 6. The particular sample thickness value (da) can also be retrieved by the application from a data file such as the one presented in table 2.4, by pressing the button to the left of the text box.

Since two fluxes exist, the diffusion time of water in a tissue sample can also be acquired. By selecting the time at the highest concentration of OCA.

Once again, the diffusion time value of water (tw in figure 2.13) and sample thickness (dw in figure 2.13) obtained at that particular time in the correspondent treatment must be introduced manually in the lower left side of the window presented in figure 2.13. The desired thickness value (dw) can also to be retrieved from a graph that the application generates once the thickness data file (such as table 2.4) is uploaded. This file contains the sample's thickness as a function of time at highest concentration. Since the diffusion time of water and the respective thickness value are now known the diffusion coefficient of water is also calculated (equation 6).

The calculation of the diffusion coefficients is made, for water and for OCA, either with loadable thickness files or simply by inserting the values of thickness and diffusion time manually. Such procedure facilitates the work flow of the application and removes the need of a specific file format such as the one shown in table 2.4. The buttons to load the thickness files (in figure 2.13, to the left of the text boxes) become visible after the OCA diffusion time and water diffusion time are manually inserted. The calculated results also become visible and are updated every time one of the buttons are pressed or the times and thickness are changed manually. Figure 2.14 shows the results in the group bellow the text boxes used to introduce the diffusion time and sample thickness values.

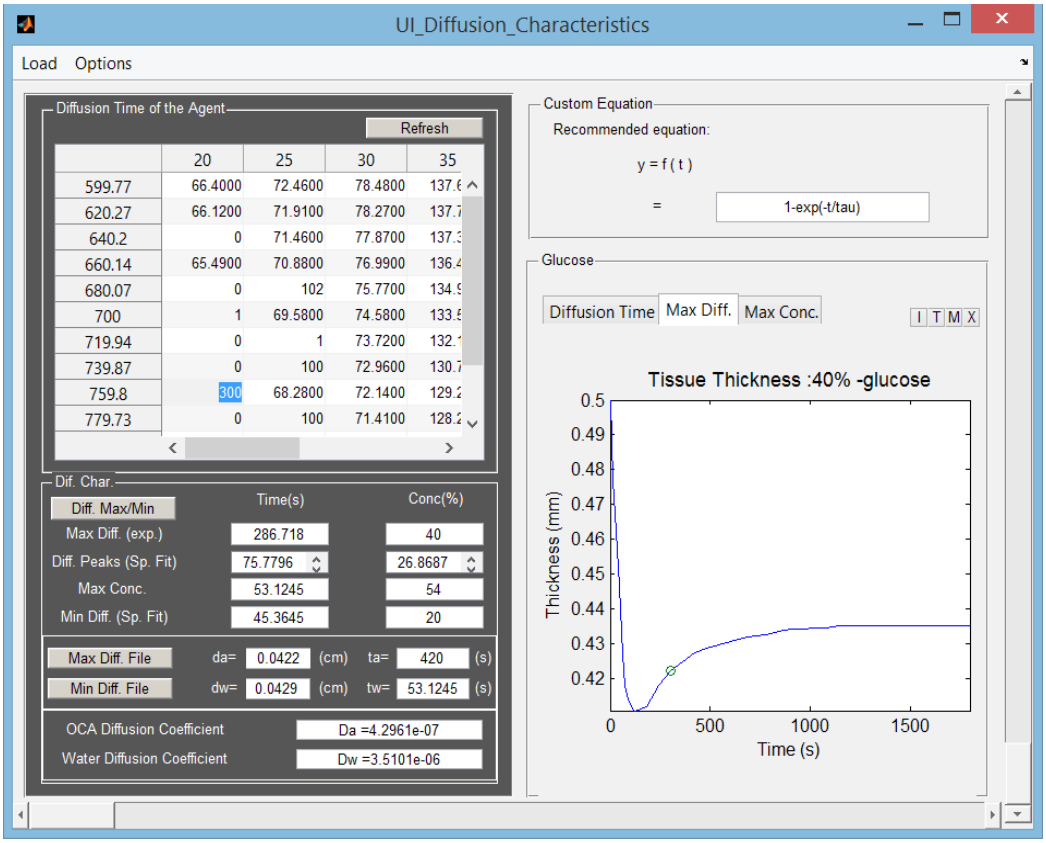

Fig. 2.14 Application Interface.

To make sense physically with the experimental data, application limits of the values introduced were set for thickness and time so they can be used in the calculation of the diffusion coefficients. For thickness, only values between 0

and 20.99 cm are accepted. For diffusion time, only values found in the table presented in figures 2.12 or 2.13 with a margin of an extra two minutes (120 s) are allowed. Symbols and special characters are not accepted. When any of these statements are not met, the value becomes 0, preventing any kind of application malfunction. The reason for these limits being so large revolves in the capability of giving the user some liberty to observe different results by modifying its parameters.

Figure 2.14 shows one of the limits stretched to its maximum range, where the diffusion time of the OCA is tested with 420 s, which corresponds to the maximum existent value, 300 s (selected blue cell from the table), plus 120 s.

The two extra tabs that appear from figure 2.13 to 2.14 means that the thickness files for maximum diffusion time (OCA) and maximum concentration (water) were loaded and their respective data is represented in a function. When the thickness files are inserted manually, no functions are presented. Once both calculations of the diffusion coefficients of OCA and water are made, or any of the variables used in equation 6 are changed the results are refreshed and appear in the lower left of the application window (figure 2.14).

2.9 Image and Table export

All functionalities that show tables, graphs or both, generate a tab so the data is always properly organized and in its respective position inside the window. Inside all tabs there are four small buttons, at north-east, giving some extra functionalities to the application. Starting from the left to the right, the first one is the button (I) that stands for image or graph, the second is (T) that stands for table, third button is (M) that stands for maximize and the last one (X) stands for close since its the usual symbol for such action.

The image button (I) saves all graphs shown in the concentration tabs to a folder. For example, figure 2.9 shows the maxima found of all λ in each concentration tab, by pressing the (I) button those eight graphics (20% to

54%), being the time dependence of Tc: 25%-glucose graphic shown in figure 2.9 one of those eight, are saved to a distinct folder. Depending on the tab-group that (I) button is presently located, the represented functions, separated by tabs, will be saved into different folders. Each saved image has a distinct name in order to be distinguishable between different concentrations. As well as the saved images are grouped in different folders depending on the functionality they were recorded (for example MaxDetect folder contains images that correspond to maximum detection functionality). The functionalities that use the (I) button as described above are the initial data loaded, the bandwidth generation, the maximum detection, the normalized data, the diffusion times and the thickness files.

The table button (T) also saves the tables of all concentrations, separated by tabs, to a particular folder using a very similar approach to the one used in the images save. The functionalities that can export or save the tables are the initial data loaded, the bandwidth generation, the maxima detection, the normalized data and the diffusion times (this last one only has one table, but the position of the (T) button remains in the same position, inside the tab-group, as figure 2.14 demonstrates). This means that every time the (T) button is pressed in a given tab-group, all the content from the tables inside the concentration tabs are saved exactly as they appear in the application interface. An exception is verified for the first two maxima detection tab-groups (two of three buttons that offer the maximum detection/selection functionality). The first two maxima detection functionalities do not save the tables exactly as they appear in the tab-groups. These two only save the content of the tables up to their selected maximum time value. For example, in figure 2.9 the tables still have all data, ranging from 0 to last value which in this case is 1800 seconds. However, if the (T) button is pressed in this tab-group it will only save the content ranging from 0 to maximum value selected (each tab has its respective maximum shown in the small table south-west next to the Maximum label).

The tables are saved this way because they belong to the same functionality as the graph that represent T_c time dependence from 0 to maximum value (graph that neglects the data after the saturation regime). For that reason all maximum detection functionalities share the same table file. Consequently, every time these tables are saved by pressing the (T) button, the file is overwritten, avoiding any redundancy.

Two table export formats were tested during the development, *txt* and *xlsx*, where both offered conclusive outputs (meaning the data was correctly saved in terms of quality). The performance tests provided very different results. The *txt* test to export the tables took only a few seconds (not more than 10 seconds), but the *xlsx* took a few minutes to export all the tables (depending on the length of data it could go up to 10 minutes). Even if the *xlsx*, *xls* or other data-sheet format is more adequate to work with tables the application was implemented to export the tables in *txt*. Because the time to export to *txt* and then convert that file with a calculation processing software is still far less time consuming than directly export to *xlsx* format (the conversion from *txt* to another format like for example OpenOffice Calc is almost instantaneous). By taking special notice on the space delimiter, and unite the delimiters if necessary since the tab delimiter may be counted multiple times . A reduction of the exportation time of the table content to a data-sheet format was possible. Less than half the time compared with the application exportation directly to *xlsx*.

These two exportable items (images and tables) are saved inside the root of a Latex document named *Report*, allowing the user to generate the document with the help of *Tex Works*, or simply use those items (tables and graphs) in another document processing software, since the format of those items are quite common.

Third button is the (M) button. This button allows the data of the corresponding tab, and only that tab (meaning it does not display the content of the entire tab-group at the same time), to be displayed in full screen as figure 2.10 demonstrates. This maximized window uses the MATLAB figure

function to display the data in graphs. When the data corresponds to multiple wavelengths it will first display all wavelengths in the same tab. By taking a closer look to figure 2.10, the first tab is named 25%, meaning that in that tab are represented all wavelengths corresponding to the treatment with 25%-glucose. The next tabs will display its respective λ in a unique tab named after it. This option provides a higher level of detail since the functions are maximized, and it allows the user to modify graphic properties as he desires through the plot options (edit button in the figure menu). Figure 2.15 shows the plot with corresponding properties that can be edited by the user.

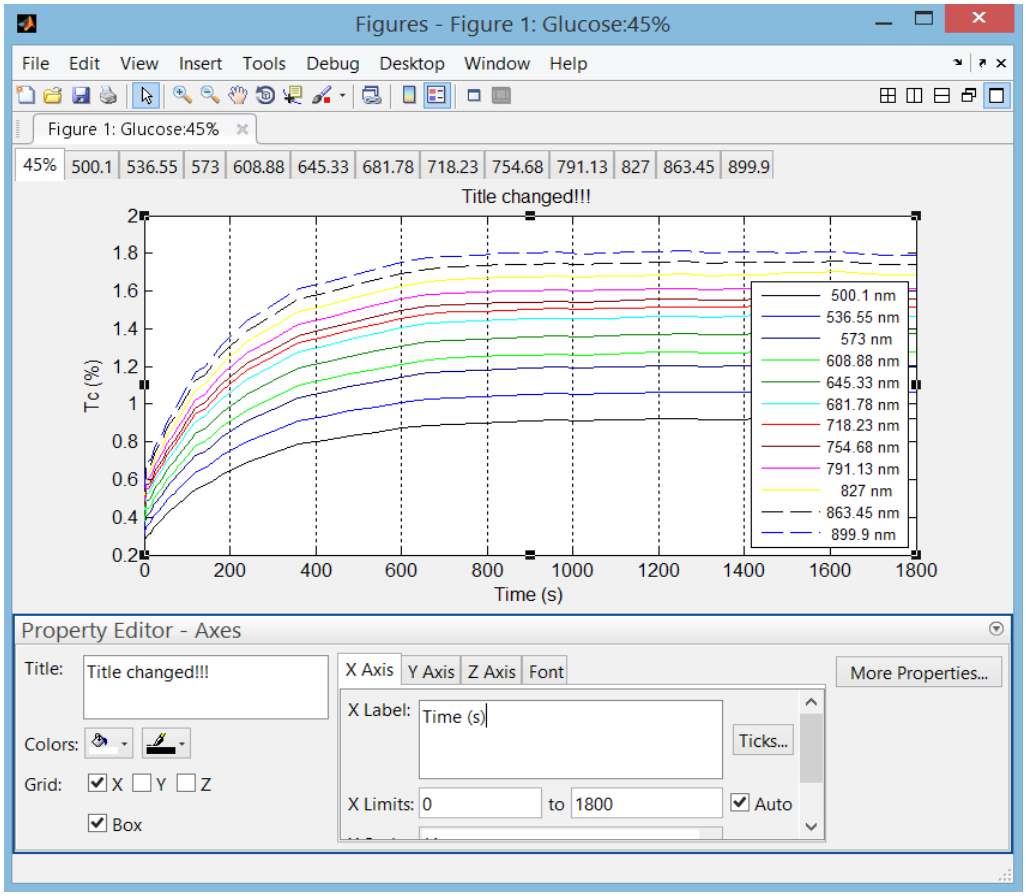

Fig. 2.15 Maximized Window. Bandwidth [500 900], 12 wavelengths.

All user interface (UI) controls inside this figure function (figure 2.15) must be in pixel units in order to prevent the occurrence of an error (warning) in the *hgconvertunits* internal function of MATLAB, which unfortunately only works

in pixels. In the application, the units remained normalized in order to keep the original concept of the maximize functionality.

The fourth and last button is the (X) button, which closes the entire tab-group, meaning it closes all related tabs to its parent. Tab-groups should be closed before running another functionality in order to prevent overlapping of groups.

2.10 Menu functionalities

The developed application has a menu to help and facilitate the working flow of the application which consists in a small bar with multiple buttons in the top of the window. Such menu can be seen in figure 2.16.

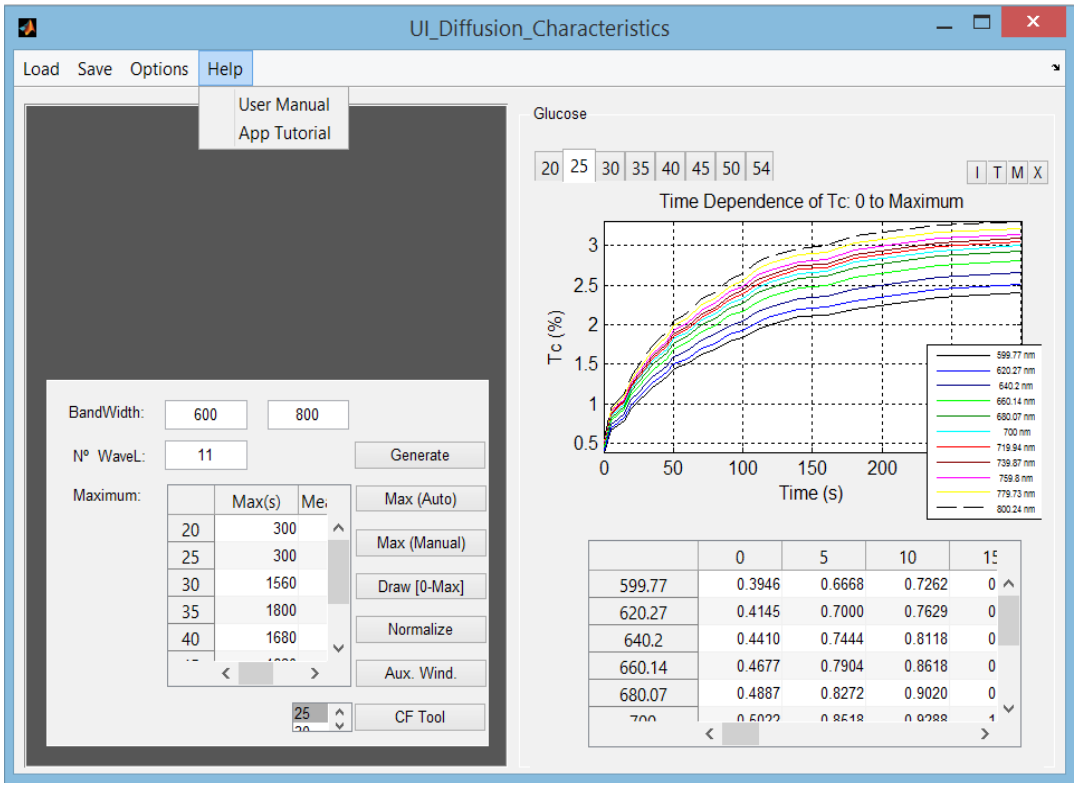

Fig. 2.16 Application Interface with Grid on and App. Log hidden. Draw [0-Max] pressed, and respective data [0-Max] seconds shown in tabs.

The help sub-menu seen in figure 2.16 allows the user to open a tutorial to highlight some common errors that can occur while using the application and

the usual properties of the initial loadable files. The other option under the help sub-menu (App Tutorial) opens a step-by-step tutorial with images explaining some of the functionalities and normal sequence of the application with simple examples.

Options sub-menu, seen in figure 2.17, allows the graphs to be drawn with a grid but it should be noticed that the grid is only applied on the following graphics, not the ones already displayed on the interface. Closing and re-opening the desired graphics will apply the grid.

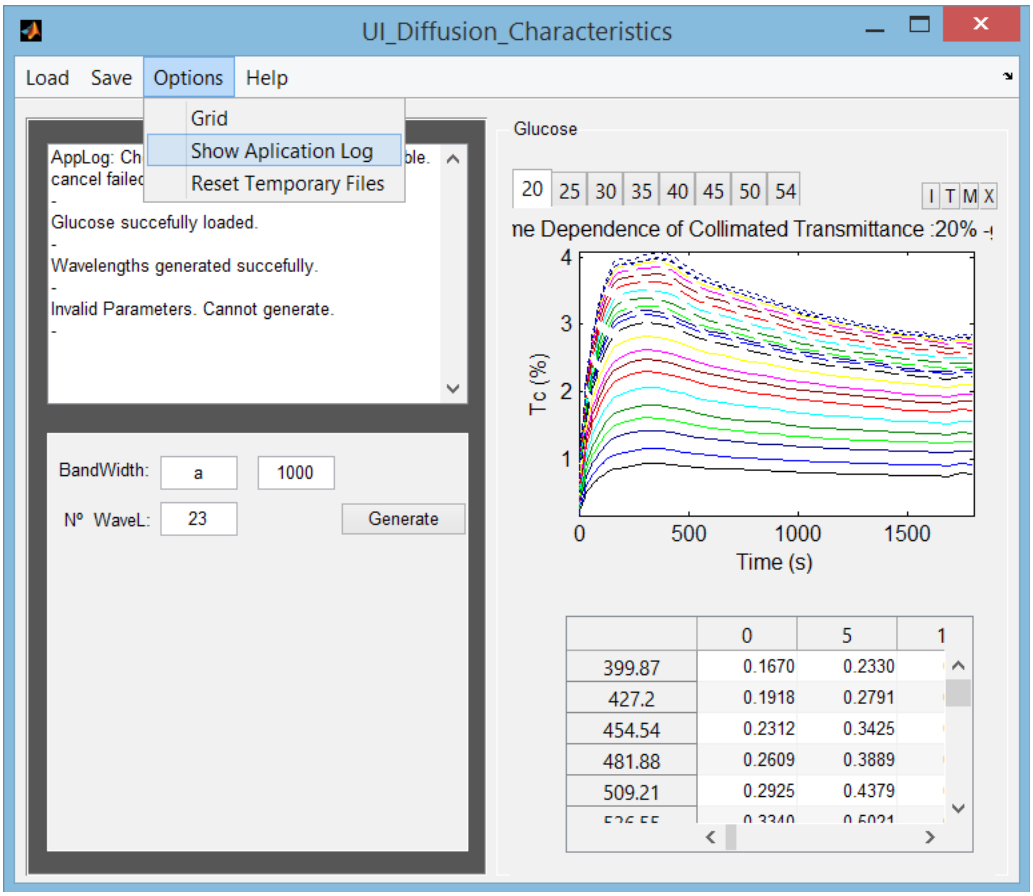

Fig. 2.17 Application Interface showing the sub-menu options and warnings when a validation is called.

Hiding the application log is also possible, the warnings will continue to be written for some of the functionalities, so the user can always look up the progress and flow of the application. This is seen in figure 2.17. The validations are called and presented in the app.

In future versions the Application log will be redesigned to return warnings for all functionalities, as well as generate a small script that will record all the application history for performance and error analysis (improvement data, feedback). Since at this stage, the intended purpose of the developed application is purely academic, the full strength of this functionality has been postpone.

Resetting the temporary folder (options sub-menu) deletes all temporary files automatically saved through the course of the application. Using this functionality implies starting all over again from the beginning. The temporary folder (file explorer), must also be closed to prevent any kind of error. Same goes for any file or folder that the application is using, since it cannot be deleted or changed if the file is being used by another program.

Saving the work progress has two options. It has an option to export the saved graphics into a pdf file, with the use of Tex Works, and saving the data for posterior use of the application. This second option can be done at two specific checkpoints. After the maximum is found and data shown is ranging from 0 to the defined maximum (last button from figure 2.9) and when the refresh button on the auxiliary window is pressed (figure 2.12).

The load sub-menu, has the option to load the initial data (figure 2.2) previously described in the beginning of this chapter, it loads new data to be processed. The load sub-menu also has a functionality to load a previously saved work, if it was correctly saved after reaching one of the two checkpoints.

The sub-menu can also run the cftool with no data automatically inserted. This allows the user to load the content of the experiments that were previously recorded and continue the fitting process. These sessions can be loaded and saved in *sfit* format (*sfit* is the type of file which sessions are saved using cftool).

3 Chapter 3

3.1 Discussion of results

As indicated in chapter 1, there are several OCAs known today with great potential to create optical clearing effects in biological tissues [1-16]. With the objective of estimating their diffusion properties in different biological tissues, the developed application can be used. This chapter presents the study of glucose diffusion into skeletal muscle and compares the results with the ones found in literature in order to validate the level of efficiency of the developed application.

3.2 Analysing initial data

The initial experimental data used in this experiment was obtained by measuring the T_c spectra using the methodology explained in chapter 2 [1-2]. Each measured spectrum contains values from 172 to 1100 nm and spectra were measured during treatments with different aqueous glucose solutions for 30 min. Due to the noise on the sides of the measured spectra, the bandwidth of the spectrometer used is clearly not suitable to perform the desired calculations.

By discarding this noise the various spectra become more reliable and trustworthy for proceeding with the calculations. Considering the acquisition band of the spectrometer and since the lateral regions of the spectra cannot be trusted, the application allows spectral data below 400 and above 1000 nm to be removed.

Such lateral noise in T_c spectra can be seen in figure 3.1.

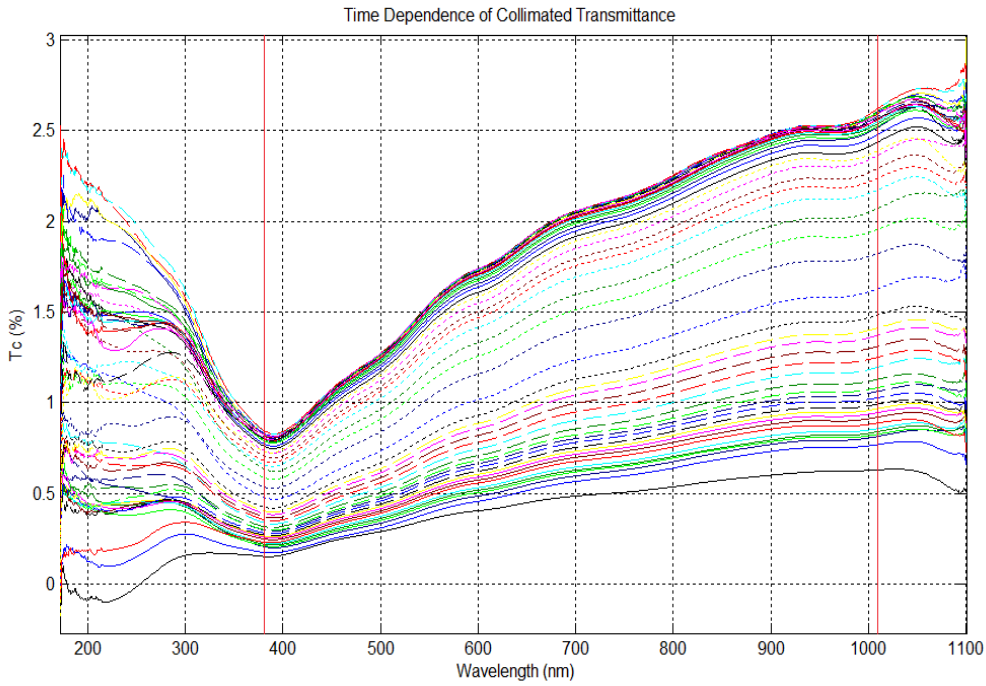

Fig. 3.1 Measured T_c data loaded showing high level of noise at the sides, and consequently not trustworthy. This graphic represents 40%-glucose.

This selection prevents unnecessary data to be used in the following steps, minimizing memory resources in the computer. If the spectrometer range is properly calibrated before the measurements, this step could be avoided, increasing the efficiency of the initial upload.

It is possible to obtain the slit through the wavelengths spacing. Table 3.1 demonstrates the wavelengths of the measuring spectrometer, separated by 0.57 nm.

Table 3.1 Wavelengths configured on the measuring spectrometer.

Wavelengths									
171.49	*172.06*	*172.63*	*173.2*	*173.77*	*174.34*	*174.9*	*175.47*	*...*	*1100.4*

This λ spacing of 0.57 nm is not the best option to display numbers rounded to units. Since the desired bandwidth for calculations presented in literature

is from 600 to 800 nm and the λ step is 20 nm, then it is safe to claim that the slit should be 0.5 nm or 1 nm (closest values to the left and right of 0.57 nm). This way the selection of wavelengths from the bandwidth would correspond to the theoretical values, and eliminate the induced error.

3.3 Highest scattering band

To study OCA diffusion in a biological tissue, it is necessary to consider a bandwidth where light scattering dominates strongly over light absorption [1, 26].

For skeletal muscle light scattering domination is between 600 and 800 nm [1-2]. On the other hand, each biological tissue has a characteristic T_c spectral form that must be previously known in order to perform a correct identification of the desired bandwidth in the application. For fibrous tissues like the skeletal muscle, it is expected an increasing behaviour of T_c with λ in the area of scattering domination [32].

The samples used in the experiment were collected from the abdominal wall muscle block of a Wistar Han (rat), and according to literature the band where this type of tissue is defined by scattering ranges from 600 nm to 800 nm, and is presented with a 20 nm spacing [1-2].

There are a total of eight OCA concentrations that were considered in this study. All treatments were performed for 30 min.

Figure 3.2 presents the measured T_c time dependencies for the first four, 20%, 25%, 30% and 35%. These graphs presented in figure 3.2 have a bandwidth from 600 to 800 nm with 20 nm spacing, where the tissue spectral profile is mostly defined by scattering.

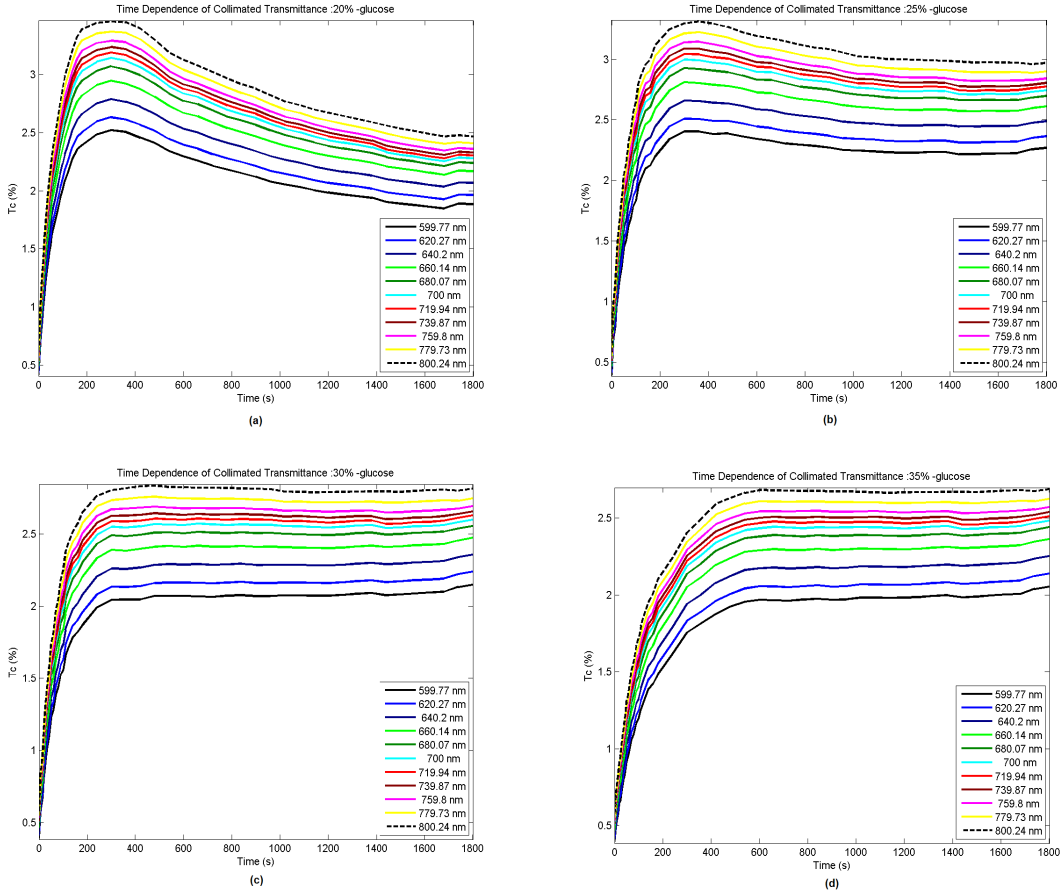

Fig. 3.2 Each graphic corresponds to a treatment with a particular concentration: (a) 20%, (b) 25%, (c) 30%, (d) 35%.

Analysing every graph individually, or by OCA concentration used in the treatment, the wavelengths maintain a very similar form, rising very fast at the beginning, before tissue saturation, and then stabilizing or decreasing, depending on the OCA concentration used. For a particular OCA concentration, higher wavelengths reach higher levels of T_c at the end of the treatment.

This fact is related to the increasing spectral form of T_c with λ seen in natural tissue [1]. Figure 3.3 presents such natural T_c spectrum. In that figure can be seen the increasing T_c with λ between 600 and 800 nm as proof of the scattering dominating behaviour in the skeletal muscle.

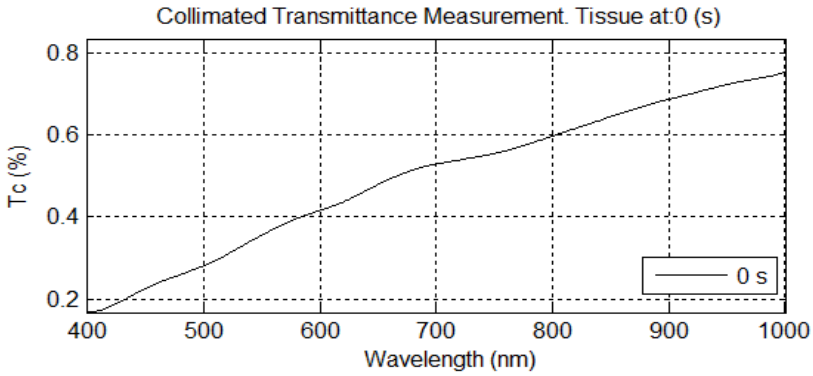

Fig. 3.3 T_c spectrum of the natural tissue. Bandwidth [400-1000].

For 20%-glucose (figure 3.2(a)) the time required to achieve saturation is short and T_c starts to decreases after reaching it. The tissue contains much less water than the immersing solution and the creation of a water flux into the tissue becomes a possibility. This means that the OCA diffusion is quite fast due to the creation of the water flux. After the initial T_c increase has ended (beginning of the saturation regime), the tissue becomes dehydrated because of the impact of outside glucose on it, and in order to restore the balance the tissue will receive water from the immersion solution which causes T_c to decrease over time until the end of the treatment. During the exchange of the contents from the tissue to the immersion solution and vice versa, the thickness of the tissue may vary over time. Such sample thickness variation has to do with the fluxes of water going out and OCA going in during optical clearing as well as difference between the molecular size of water and glucose molecules.

The same thing happens with the treatment with 25%-glucose, the initial T_c increase is fast and presents identical magnitude values to the case of 20%-glucose treatment. When it reaches the saturation regime T_c starts to decrease due to the dehydrated state of the tissue, however after saturation the decrease is not as strong as in the treatment with 20%-glucose. The increase of glucose concentration in the immersing solution gave some stability to the graph form after the maximum is reached, meaning that the

water content in solution is now closer to the free water content in the tissue.

For the treatments with 30%-glucose and 35%-glucose the diffusion is still fast, but not as fast as in the treatments with smaller glucose concentrations. Initial T_c increase seen for the treatment with 35%-glucose takes more time than in the treatment with 30%-glucose. This increase of glucose in the immersing solution indicates that glucose takes more time to diffuse into the muscle and saturation regime is achieved later. Such behaviour is evidence that the water content in the immersing solution is approximating the water content in the tissue.

Figure 3.4 represents the experimental data collected from the treatments with the remaining four OCA concentrations (40%, 45%, 50% and 54%).

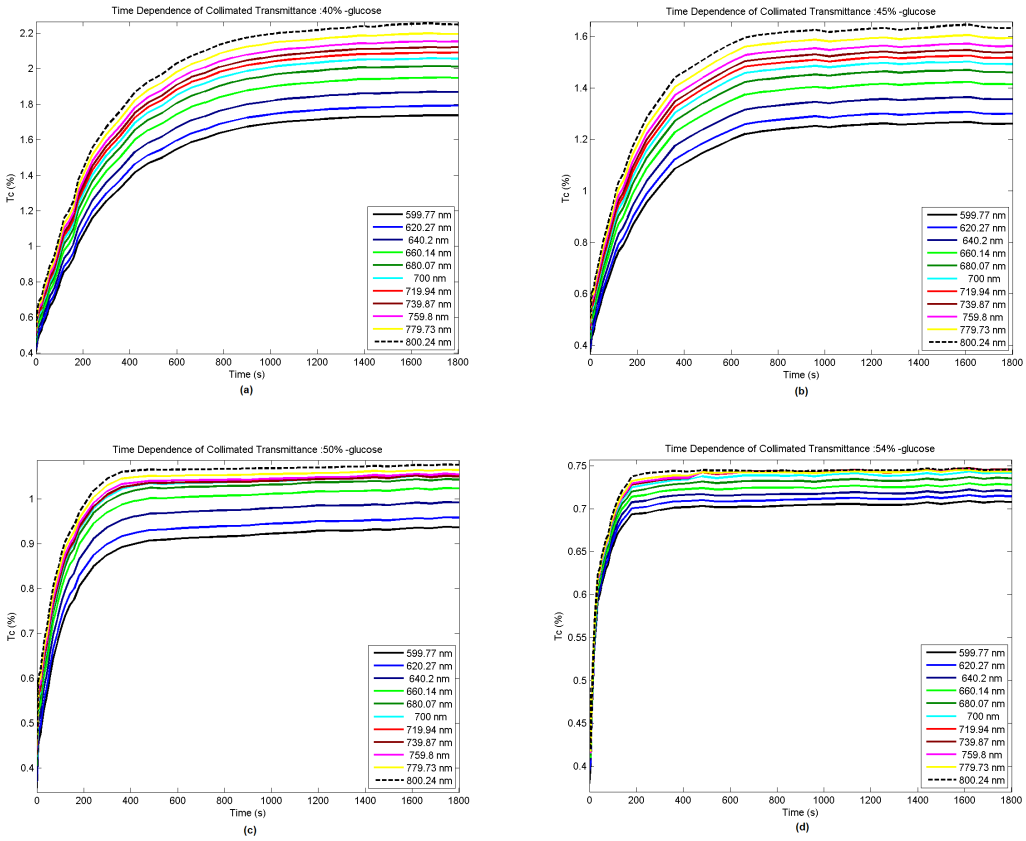

Fig. 3.4 Each graphic corresponds to a treatment with a particular concentration: (a) 40%, (b) 45%, (c) 50%, (d) 54%.

These treatments, from figure 3.4, also have a bandwidth ranging from 600 to 800 nm with 20 nm spacing, where the tissue spectral profile is mostly defined by scattering.

When the water content in the immersing solution is exactly the same as the free water content in the muscle, no water flux is verified, in or out of the tissue. Such water equilibrium maximizes glucose diffusion into the muscle. This happens when the OCA concentration in solution is close to 40% (figure 3.4(a)) and corresponds to the optimized OCA diffusion time.

For glucose concentration in solution higher than 40%, there is a gradual decrease in the highest T_c values as the sequence of graphics demonstrate from 2.2% to 0.75%, for 40% to 54% (graphs *a* to *d* in figure 3.4). The most important aspect to retain from this sequence is that the diffusion time appears to decrease (fast diffusion), because water flux is no longer zero but this time water is moving out from the tissue to the solution due to the excess of glucose in solution (glucose hyper-osmolarity). The time response seams much shorter but in truth there is no inversion after the initial fast rise, so the diffusion time is very fast in the beginning, and takes the rest of the time to reach its maximum.

3.4 Preparing data for diffusion characteristics calculation

To calculate glucose diffusion coefficient through equation 6 the real diffusion time of glucose and sample thickness for that particular treatment and particular time are required. Time dependence measurements of the tissue thickness were performed from similar samples under treatment with all eight immersion solutions (20% to 54%), so the thickness values could be detected at the desired time of a particular treatment.

The beginning of the saturation regime for each treatment was detected and can be seen in table 3.2.

Table 3.2 Beginning of saturation regime for each treatment with different glucose concentrations in the immersing solution.

Glucose Con.	*20%*	*25%*	*30%*	*35%*	*40%*	*45%*	*50%*	*54%*
Time (s)	*300*	*360*	*480*	*660*	*1800*	*1800*	*1800*	*1800*

The remaining T_c time dependence (after the beginning of the saturation regime) was neglected in each case. The entire time dependence delimited by the beginning of the saturation regime was vertically displaced to have $T_c=0$ at $t=0$ (natural tissue). After vertical displacement, a normalization of the entire time dependence to the highest value was made to obtain the entire time dependence between 0 and 1. Such procedure is used so that the corrected time dependence can be fitted with appropriate equation (5), in order to obtain the diffusion time of each λ with a particular solution.

3.4.1 Step 1 – Maximum Selection

As expected the beginning of the saturation regime occurs early in treatments with low and high glucose concentrations and take more time to occur for treatments with intermediate glucose concentrations, revealing the importance of the osmotic pressure of the agent in the immersing solution.

Figure 3.5 presents the T_c time dependencies obtained from the treatments with four solutions with smaller glucose concentrations for a time scale delimited by the beginning of the saturation regime. Comparing between graphs in figure 3.5, they now appear to share a very similar form, since all the remaining time dependence after the beginning of the saturation regime has been removed.

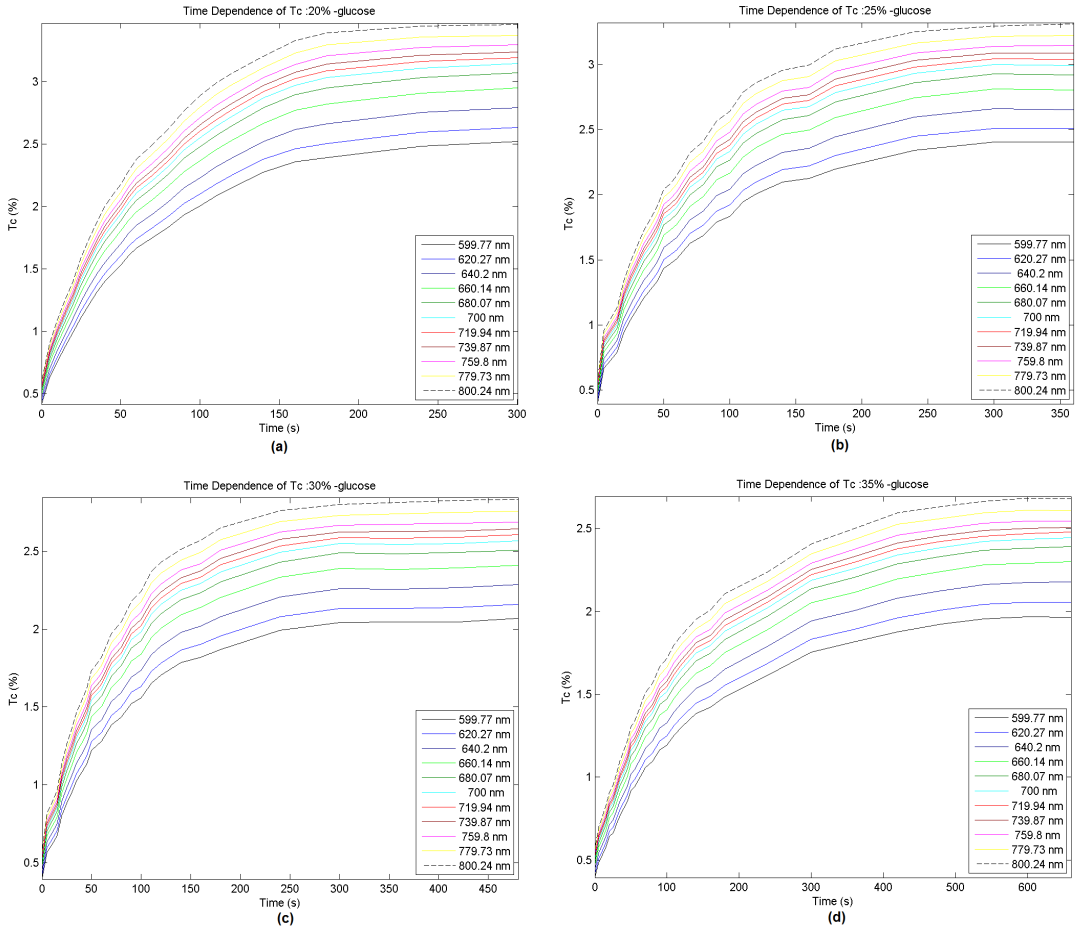

Fig. 3.5 T_c time dependence for wavelengths between 600 and 800 nm (with 20 nm spacing) for the treatments with glucose concentrations: (a) 20%, (b) 25%, (c) 30% and (d) 35%.

The T_c time dependencies for the remaining treatments (with glucose concentrations in solution from 40% to 54%) are presented in figure 3.6. With the exception of the treatments with higher concentrations (graphs (c) and (d) in figure 3.6), they also share the same form. The data for the two treatments with the highest glucose concentrations ((c) 50% and (d) 54%) were not trimmed since in these cases the T_c values continued to increase until the end of the treatment (1800 s or 30 min). Although such increase is seen during all treatment, the saturation regime is linearly increasing for these treatments and the mean slope of that linear increase is low.

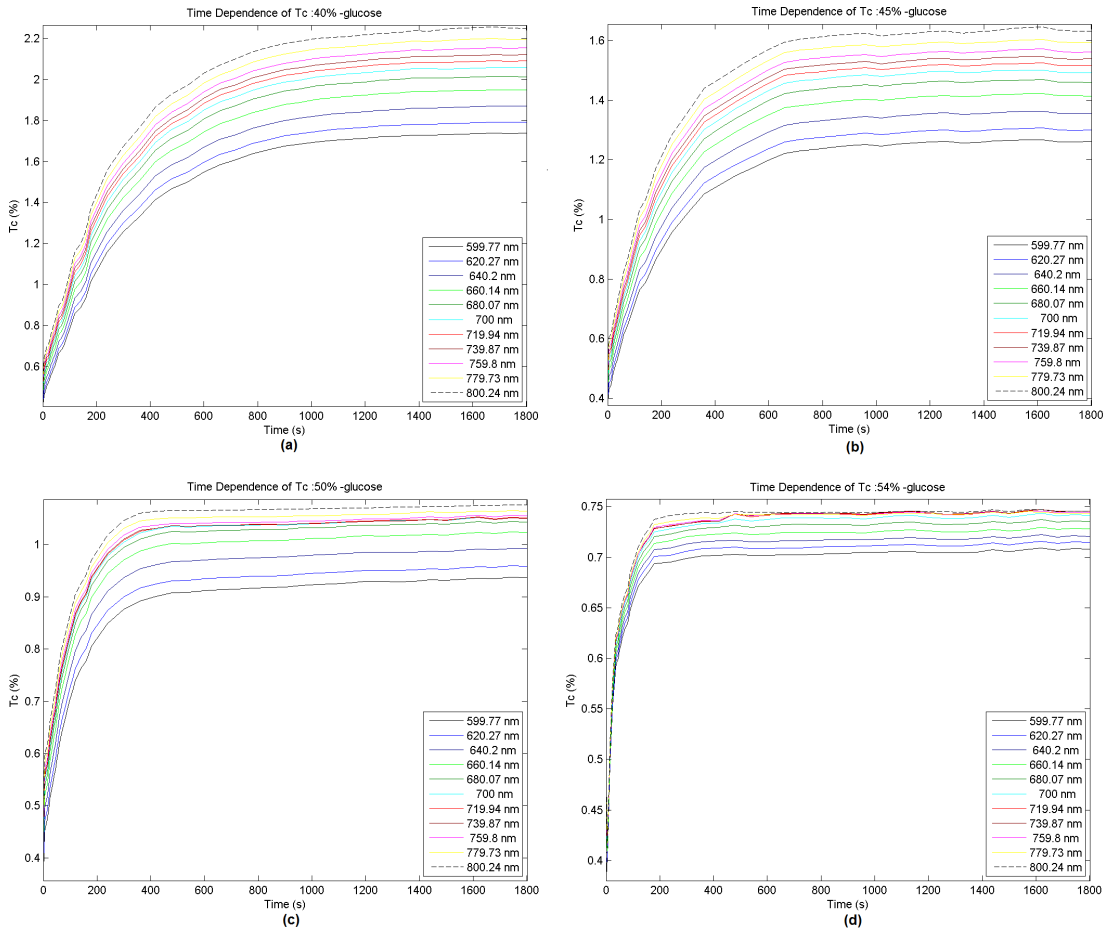

Fig. 3.6 T_c time dependence for wavelengths between 600 and 800 nm (with 20nm spacing) for the treatments with glucose concentrations: (a) 40%, (b) 45%, (c) 50% and (d) 54%.

The two highest concentrations in the immersing solution (50% and 54% from figure 3.6) may indicate that glucose continues to diffuse into the muscle, but at the expense of some tissue swelling. Since this tissue swelling is not very high (small slope in linear saturation), the entire T_c time dependence can be considered in our calculations without leading to a false determination of the diffusion time for these cases.

Evaluating the treatments with low and high glucose concentrations revealed the high importance of the osmotic pressure of the agent in the immersing solution. For such concentrations the outside osmotic stress forces water to

leave the tissue at early treatment (observed in treatments with 20%-glucose, 25%-glucose, 50%-glucose and 54%-glucose). The difference is that for low glucose concentrations there is a mixed flux with water flowing out and glucose flowing in. For highly concentrated solutions the flux that dominates is the water flux out of the tissue. Ideally, with an over saturated glucose solution the corresponding treatment would stimulate only water dehydration. For that case no glucose diffusion would occur in the 1800 s period. This is already observed in the treatment with 54%-glucose.

In opposition, for intermediate glucose concentrated solutions, longer diffusion time values are seen. This means that the tissue dehydration mechanism does not dominate optical clearing operations. Instead, the RI matching mechanism dominates optical clearing, since glucose takes more time to flow into the muscle as a consequence of similar water contents in the solution and in the tissue (free water part). One of the purposes of the application developed is exactly to determine the ideal OCA concentration in solution, which is associated with maximum OCA diffusion and no water flux. Only for that particular concentration the necessary means exist to calculate the real OCA diffusion time.

3.4.2 Step 2 - Normalization and tissue darkening at high concentration

As it was mentioned above, one of the procedures to set the data correspondent to each λ within a treatment ready for fitting is to displace it vertically so that for natural tissue there is 0% T_c value. After the vertical displacement, the entire time dependence is normalized to the highest value previously obtained at the beginning of the saturation regime. After the normalization procedure, all the T_c time dependencies that correspond to all treatments (with the exception of the treatment with 54%-glucose) have the expected increasing behaviour over time.

For the treatment with 54%-glucose such behaviour also occurs, but with a small T_c decrease observed within the first 10 s. This initial T_c decrease indicates some tissue darkening that occurs due to the high glucose concentration in the immersing solution. In this case, the osmotic pressure over the tissue is very high and leads to a strong and fast tissue dehydration. As the tissue loses water, the muscle fibers approach each-other and sample thickness decreases also. These variations turn the tissue more turbid at this early stage of treatment. Only a few seconds after, as glucose begins to enter the interstitial fluid of the muscle and forces the muscle fibers to separate, T_c rises. Such variations indicate that in the first few seconds of this particular treatment no glucose flux occurs. After this first seconds, glucose flux into the muscle is also very small, since for this treatment is possible to observe the beginning of the saturation regime very early, within the first minute. The initial decrease in T_c values for this treatment can be seen in figure 3.7 and table 3.3.

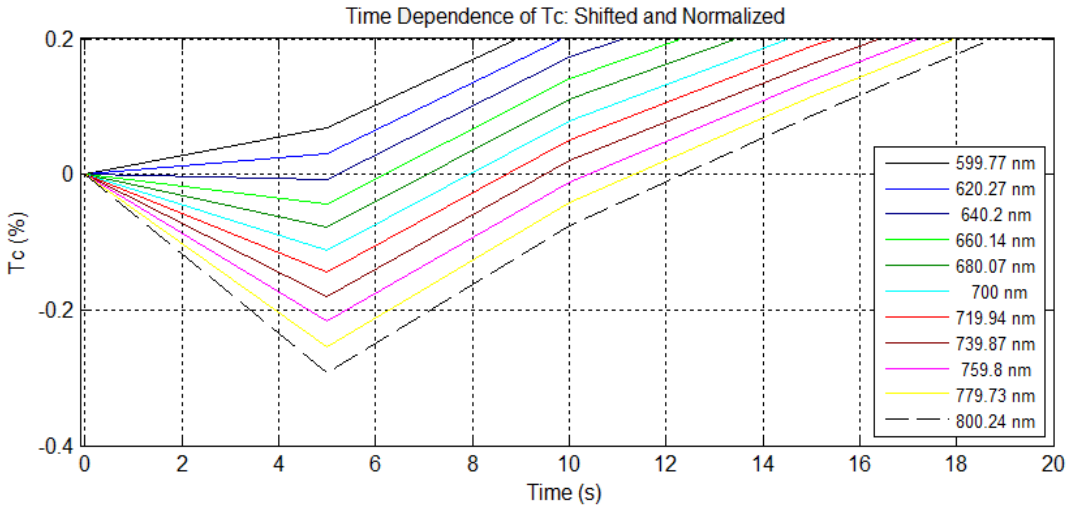

Fig. 3.7 First 20 seconds of the T_c time dependence for all wavelengths with 54%-glucose after vertical displacement and normalization.

Table 3.3 Excerpt data from 54%-glucose experiment, vertically displaced and normalized .

		Time (s)				
		0	5	10	15	20
λ (nm)	599.77	0	0.0694	0.2373	0.3559	0.4589
	620.27	0	0.0311	0.2051	0.3271	0.4336
	640.20	0	-0.0087	0.1726	0.2981	0.4087
	660.14	0	-0.0445	0.1406	0.2698	0.3845
	680.07	0	-0.0782	0.1098	0.2414	0.3609
	700	0	-0.1117	0.0795	0.2137	0.3381
	719.94	0	-0.1447	0.0510	0.1885	0.3181
	739.87	0	-0.1808	0.0201	0.1628	0.2980
	759.80	0	-0.2170	-0.0115	0.1383	0.2791
	779.73	0	-0.2543	-0.0424	0.1138	0.2597
	800.24	0	-0.2936	-0.0762	0.0876	0.2387

Nevertheless a T_c decrease can be seen for some wavelengths in the first 5 s of treatment, such behaviour does not interfere in calculations. The negative values presented in table 3.3 are not real, since T_c has always positive values. They only take negative values due to T_c vertical displacement procedure. When performing the fittings for each λ datasets, all experimental data points are considered and only at 5 s it can be seen a small decrease for some wavelengths. The fitted curve will have increasing behaviour in all cases, so this tissue darkening is somehow non significant and though implicitly neglected in calculations procedure.

3.4.3 Curve fitting

3.4.3.1 Confrontation of calculated results with literature data

With the use of curve fitting tool the following tables were obtained (table 3.4 and 3.5). The data in these tables represent the diffusion times of the net flux induced by each particular treatment. The cftool is a MATLAB tool box, with multiple options to fit the curves with equation 5, thought some small differences are to be expected.

Table 3.4 Diffusion time values obtained through curve fitting tool.

		Glucose Concentrations (%)							
		20	25	30	35	40	45	50	54
λ (nm)	599.77	68.0	74.6	80.3	142.2	310.1	219.1	112.7	51.3
	620.27	67.7	74.6	80.2	142.7	309.2	217.0	109.7	52.8
	640.20	67.4	74.0	80.2	142.5	308.5	214.3	108.2	53.3
	660.14	67.1	73.5	78.9	141.9	306.2	213.3	107.6	55.8
	680.07	66.4	72.7	78.2	140.2	304.4	210.9	103.2	57.0
	700.00	65.6	72.0	77.0	139.1	300.7	208.9	102.6	58.1
	719.94	65.0	61.3	76.1	137.7	298.1	207.6	101.4	59.0
	739.87	64.6	71.1	76.0	136.1	295.9	207.1	100.9	61.6
	759.80	65.2	70.3	75.3	135.3	293.0	206.3	99.2	63.2
	779.73	64.3	70.0	74.4	133.9	288.9	201.4	97.2	64.5
	800.24	63.6	68.8	73.8	129.7	285.0	197.5	96.9	65.8

The mean diffusion time values for each particular treatment displayed in table 3.5 were calculated with the diffusion times obtained through the cftool and saved in table 3.4.

Table 3.5 Mean and standard deviation of the diffusion times (data calculated with the application).

Data calculated through the developed application								
Conc (%)	20	25	30	35	40	45	50	54
Mean (s)	65.9	72.1	77.3	138.3	300.0	209.4	103.6	58.4
S.D. (s)	1.5	2.0	2.4	4.2	8.6	6.5	5.3	4.9

Figure 3.8 represents the diffusion time of glucose as a function of the concentration in solution (glucose).

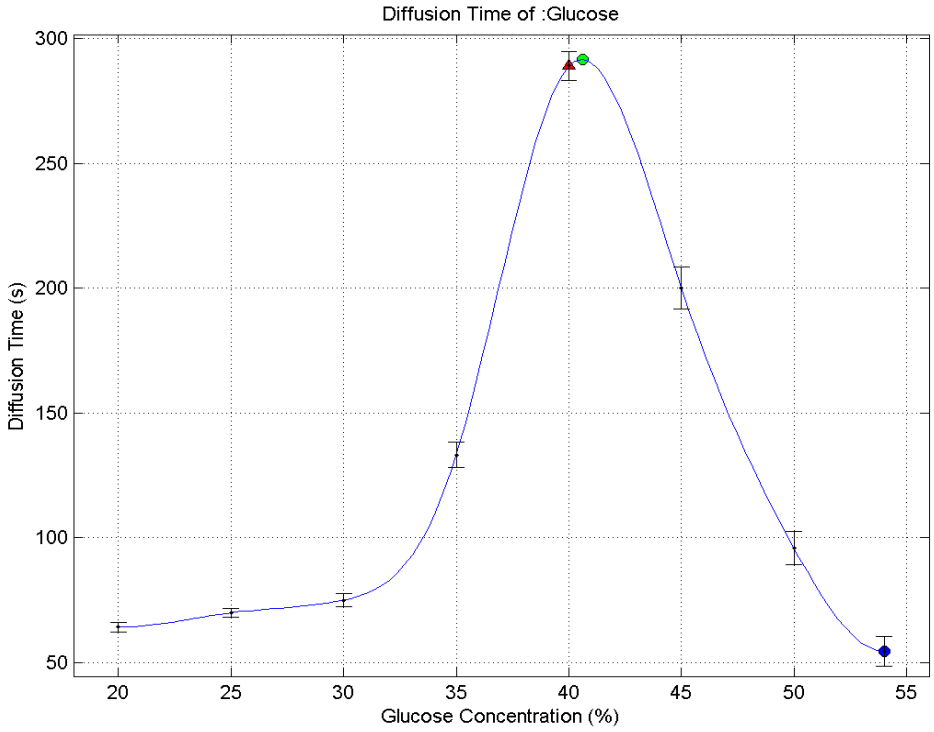

Fig. 3.8 Diffusion time of glucose as a function of glucose concentration in solution.

In figure 3.8 the red triangle indicates the maximum measured time, the green dot corresponds to the maximum time detected in the spline function, and the blue dot indicates the maximum concentration.

By analysing these results there is a difference between the maximum experimental diffusion time and the maximum diffusion time detected in the spline function of 2.8 s. Considering the entire duration of the measurements (1800 s) it is not much, but the difference is important. This reflects a certain care from part of the investigator who chose the concentrations of glucose for the experiments (at least one concentration near the expected maximum). That is why the concentrations difference between the spline and measured values is also low. The glucose concentration in the solution at the maximum time in the spline is 40.6% of glucose which gives a difference of 0.6% from the nearest concentration used in the performed measurement which is 40%. The minimum diffusion time and the highest concentration are found at the same given time and concentration, which is 54% of glucose in the immersing solution at 58.4 s. Figure 3.9 shows the UI from where all the information displayed was exported or consulted.

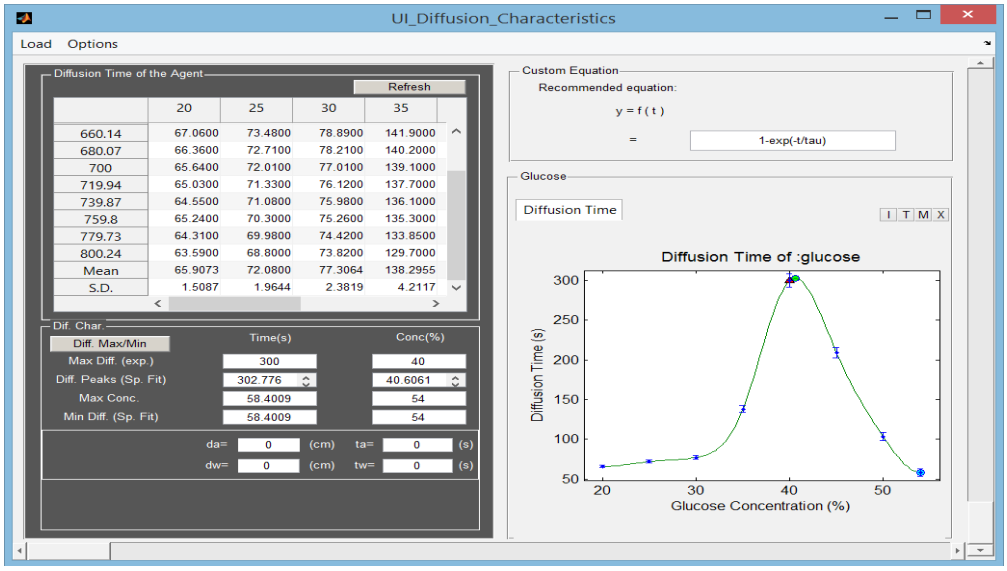

Fig. 3.9 Detection of maximum diffusion time, and concentration.

The diffusion times, when compared with the results found in the literature[1-2] coincide (unity values) and for this reason the functions are also similar in form. Which proves that the application is able to reach the same results but in a much more efficient way. In literature [1-2] the maximum diffusion time was reported as 302.9 s (maximum from the plot) at 40.5%-glucose. The minimum diffusion time and maximum concentration are also found at the same given time, 58.4 s, and concentration, 54% of glucose. The following table 3.6 indicates the mean diffusion times and respective S.D., collected from literature, using the same equation (5).

Table 3.6 Diffusion times for glucose solutions, found in literature.

Data collected from literature								
Conc (%)	*20*	*25*	*30*	*35*	*40*	*45*	*50*	*54*
Mean (s)	65.9	72	77.3	138.3	300	209.4	103.6	58.4
S.D. (s)	1.8	2	2.2	4.3	4.9	7.4	7	7.9

Comparing between calculated data on table 3.5 and literature data on table 3.6 we see that the mean diffusion time values are almost the same. On the other hand the S.D. values are a little different.

The fact that the values are different does not invalidate the accuracy of any of the diffusion times. Instead, it shows that the diffusion time values presented in literature [1-2], that were obtained through cftool (MATLAB, version 2009b), are not exactly the same as the values obtained with the developed application (MATLAB, version 2013b). Not only different versions of MATLAB can cause this small difference, but also different options used in cftool can contribute to these results.

Table 3.7 shows the absolute difference between the results from literature and the results obtained with the developed application.

Table 3.7 Difference between the values of table 3.6 and 3.5.

Difference of both								
Conc(%)	20	25	30	35	40	45	50	54
Mean(ms)	7.3	80.0	6.4	4.5	0.0	0.0	0.0	0.9
S.D.(ms)	291.3	35.6	181.9	88.3	3672.2	940.0	1744.7	2979.3

3.4.3.2 Goodness of the fit

Comparing between both versions of MATLAB there is a slight change on the available methods, as well as the options itself may have been internally modified inside MATLAB. The parameters from the goodness of the fit, used as guide lines, are the same according to the MATLAB documentation, even if they may be displayed differently (appear in different places in the interface) [35]. In MATLAB version 2013b the goodness of the fit parameters are shown in figure 3.10.

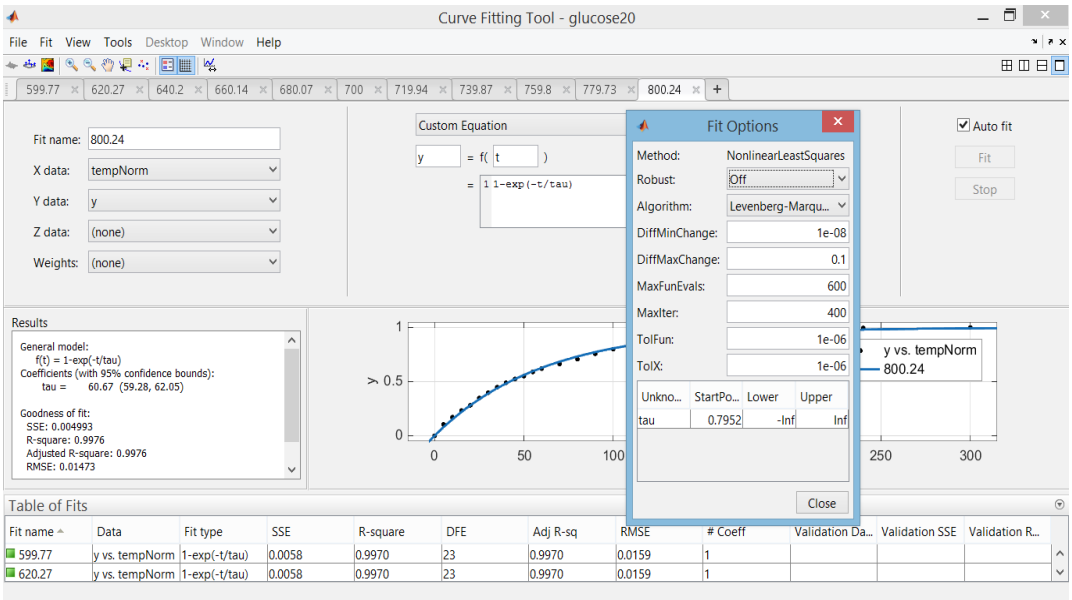

Fig. 3.10 Fit options and goodness of the fit in 2013b.

According to the detailed documentation of the cftool from MATLAB each parameter has its own weight. The SSE is the sum of squares due to error of the fit, and a value closer to zero indicates a better fit. R-square is the square of the correlation between the response values and the predicted response values. Equation 8 allows the calculation of R-square [35].

$$R^2 = 1 - \frac{SSE}{SST} \tag{8}$$

DFE is the degree of freedom in the error. Adj R-sq stands for adjusted R square, and represents the degrees of freedom, a value closer to 1 indicates a better fit. Equation 9 shows how cftool calculates its value.

$$R^2_{adj} = 1 - \left(\frac{n-1}{n-p}\right)\frac{SSE}{SST} \tag{9}$$

RMSE is the root mean squared error or standard error and a value closer to 0 indicates a fit that is more useful for prediction. Coeff stands for the number of coefficients in the model. A more detailed information about these parameters and their weight in the goodness of the fit can be found in the MATLAB documentation [35].

All of these parameters are important in a general way, but the goodness of the fit is not always ideal for the fitted curve. In the results presented from the developed application the coefficient is not taken into consideration because it is always one. Since the add of coefficients (p) do not apply here, as well as the number of observations (n), the R-square (equation 8) is equal to Adj R sq (equation 9). *SSE* is the sum of squared error. *SSR* is the sum of squared regression. *SST* is the sum of squared total. This means that by examining the SSE (the lower the better) and the Adj R-sq statistics (the

higher the better) it is possible to achieve the best fit. The RMSE (the lower the better) is useful to complement and make decisions about two similar results, since it shows the squared/standard error.

The curve fits obtained through the application for all wavelengths of each measurement (concentrations) showed in table 3.5 were calculated with 95% confidence bounds. Adj-R-sq/R-square ranging from 0.9836 to 1. SSE values were always lower than 0.07401 (and most of them were lower than 0.02). RMSE was always lower than 0.04. Only a small portion actually reached the edge of these limits, so it is safe to say that the quality of the fits, according to goodness of the fit statistics, is good.

3.4.3.3 Fit options

When comparing the variety of fit options, in opposition to the goodness of fit, in version 2009b and 2013b they are not the same. In 2009b version there were more algorithms that could be selected, such as the gaussian algorithm, but less fitting parameters than in the 2013b version. Even the cftool window, in the tests made in 2009b version, didn't work as expected in the loading of the files, because at that time it did not had tabs implemented in the internal MATLAB tool box. So the confrontation of results between the two MATLAB versions, using the developed application could not be done. In addition, due to the improvements from the 2009b to the 2013b versions, the application would have to be altered in order to work in the older version.

Figure 3.10 above shows the fit options available in version 2013b used to fit each curve. As a reminder, each cftool window like figure 3.10 demonstrates, represents one treatment (20%-glucose in this example). All the wavelengths are also separated by tabs (11 wavelengths ranging from 600 to 800 nm, with a 20 nm spacing).

Depending on which function is selected in the combo box (drop-down selector) the fit options that appear will vary. The options that are of interest to this document are the ones shown in figure 3.10, which are related to

custom equation, as a non linear least squares fitting [35]. There are 11 parameters (options) available which can influence the fit of the curve. Robust least squares fitting model can be turned off; use bi-square weights that minimize a weighted sum of squares (points too far from the line get zero weight, closer to the line get more weight); or use LAR that fits by minimizing the absolute difference of the residuals rather than the squared differences (extreme values have lesser influence on the fit). Another parameter is the algorithm used for fitting procedure which can be either trust-region (it is the default and uses the lower and upper constraints) or the levenberg-marquardt when the fit is not reasonable (does not use the upper and lower constraints). The coefficient starting point, also influence the results and must be selected according to data used as well as the upper and lower constrains of the fitted coefficients. The finite differencing parameters (DiffMinChange and DiffMaxChange) changes the limits in coefficients for finite difference Jacobian. The MaxFunEvals defines the maximum number of function models that are allowed to evaluate. MaxIter is the maximum number of fit iterations. TolFun is the termination tolerance, stopping conditions involving the model of the function. TolX is also a termination tolerance but involving the coefficients.

Most of these parameters remained in default, since they are mainly affected by the number of coefficients (in the model used there is only one coefficient, which can be confirmed in figure 3.10 in the south, at table of fits). Starting point was one of the parameters (using coefficients) that needed a particular attention since it takes a random value in the beginning of the fit (because it is a custom non-linear model [35]). As well as the lower and upper constraints that allow the definition of the tau boundaries. The termination tolerances were also left in default.

As a resume, during the fit of the curves the parameters that have been changed more often were the robust least squares, the algorithm used, the starting point, and the lower constraint. This demonstrates that there are so many parameters that can be changed, that the diffusion times (results) may

vary depending on the weight given to each one, by the user. It is also worth mentioning that the identification of the beginning of the saturation regime can also influence these results since it will vary the initial data used to fit the curves.

3.5 Diffusion characteristics

Diffusion characteristics of water and glucose in the tissue describe the time dependence of the two optical clearing mechanisms: tissue dehydration and RI matching.

Using the estimated mean diffusion time values from each treatment (table 3.5) a function on the glucose concentration in solution was presented (figure 3.8) so the maximum diffusion time of glucose, 302.8 s, could be determined. The corresponding concentration of glucose for this time is 40.6%. Figure 3.11 is the developed application window at its final state where the data collected and stored is returned as information.

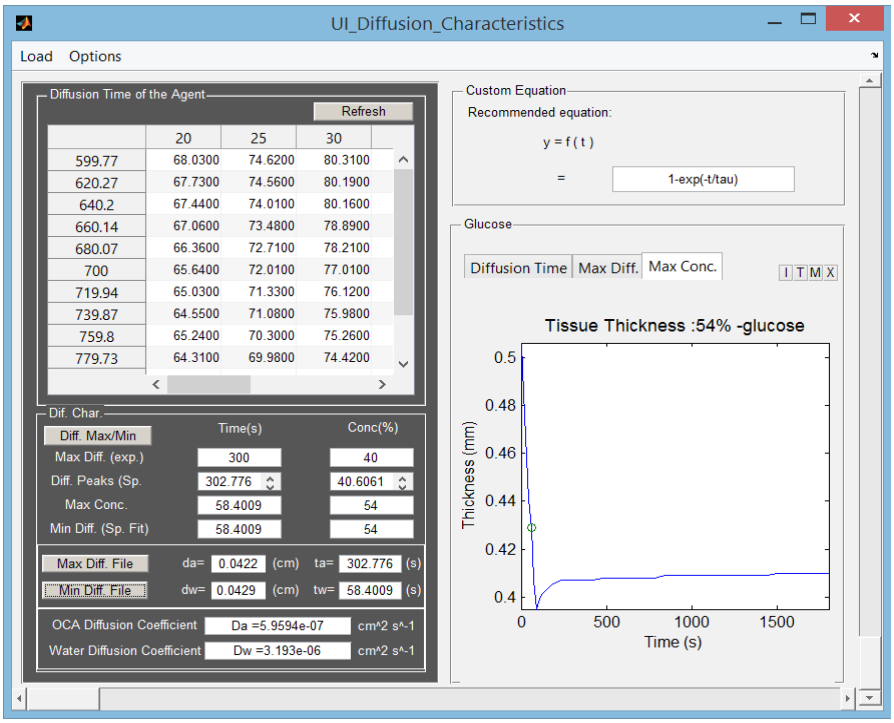

Fig. 3.11 Diffusion characteristics of glucose.

With equation 6 the diffusion coefficient of glucose could be calculated. The value obtained was $5.9594*10^{-7} cm^2 s^{-1}$.

The diffusion time of water can also be retrieved from figure 3.8. This value is 58.4 s and corresponds to the treatment with 54%-glucose.

The diffusion coefficient of water was also calculated using equation 6. The value obtained was $3.193*10^{-6} cm^2 s^{-1}$.

3.6 Flux discrimination

According to literature this type of rat tissue (skeletal muscle) has 24.4% proteins and 75.6% of water content [36]. The water in the muscle can be classified in two classes: free water and bound water [1]. Free water is the water portion in the tissue that is not connected to other tissue components and can move from one place to another or even to the outside if stimulated. Bound water is the water portion that is connected to the other tissue components and cannot move even if stimulated [37]. As indicated in this reference, bound water can even be distinguished as tightly bound and weakly bound, meaning that weakly bound water can turn into free water if the magnitude of the stimulation is high enough. For a 30 min optical clearing treatment of samples with 0.5 mm thickness, only free water moves out of the tissue, producing the tissue dehydration mechanism of optical clearing [1].

Given this information and considering the results obtained from the calculations made with the application for glucose diffusion, it was possible to estimate the free water in the rat skeletal muscle, $(100-40.6)\%=59.4\%$. Using this value with the total water content of the skeletal muscle [36], it is possible to calculate the bound water content as $(75.6-59.4)\%=16.2\%$. These values for bound and free water are almost equal to the ones reported in literature [1], with a difference of 0.1%.

The results of calculations made with the application provided the individual diffusion time and diffusion coefficient values for glucose and water in

skeletal muscle. These diffusion properties are sufficient to discriminate the two fluxes associated with the individual optical clearing mechanisms: tissue dehydration and RI matching. Some additional and complementary calculations are now possible for researchers. Since the diffusion properties here calculated are within reach, it is now possible to use equations 1 to 6 to determine the glucose concentration inside the muscle for any treatment with a particular glucose concentration, or alternatively simulate T_c time dependence for a treatment before it is studied. Such simulation calculations might be included in a future version of the application.

4 Chapter 4

4.1 Conclusion

4.1.1 Application strengths and limitations

Even thought the development of this application was mainly for academic purposes the posterior use of a compiled version of it by researchers became a possibility, turning the quick and decisive tools available. The simple sequence of functionalities allows an easy use to obtain the desired diffusion characteristics of any OCA in any known tissue. The need to constantly alter or modify manually the algorithms to answer the needs of the entire process of acquiring the diffusion characteristics have been restricted to a minimum. Decreasing the time necessary to implement the methodology without the loss of quality in the results was one of the main objectives of this application. Not only the application performance exceeded the expectations (the execution of the entire methodology has been reduced to a few hours) after countless readjustments in how to improve certain recursive processes, the results obtained also converge to the results found in literature [1-2]. These recursive methods allow the processing and calculations to be done to the entire OCA, and not one experiment (concentration) at a time. This was achieved not only internally but also shown to the user with the implementation of tab-groups, which separated the experiments. Numerous validations were also implemented in order to prevent common errors, human made errors, and posterior loss of time since it had to be fixed or remade if they were not noticed in a early stage.

After all objectives were achieved, the need for additional implementations and functionalities seamed clear and necessary. Some examples are: need for fast exportation of precessed data, options to save and load to interrupt and continue the work in different times, flow consolidation, error alert options,

higher level of liberty to the user, some shortcuts, user help options and most importantly an architecture that allows additional options to be implemented. Considering these aspects the application quickly passed from a prototype developed or an academic study to improve the determination of the diffusion characteristics, to a possible tool able to complement and answer some of the needs of researchers in the field of optical clearing.

The application limits were designed and implemented to determine the diffusion characteristics of any OCA. However, the application was only tested with one OCA (glucose) and still needs more tests so that new and different results can be once again confronted with the ones existing in literature. This will not only improve the application boundaries but will also allow the detection of possible bugs that could not be detected so far with the data used and presented in this document.

4.1.2 Accuracy of results

The application is effective as long as the initial data inserted is not corrupted. The efficiency, in question of time, is considered to be very high since the application reduced the duration of the entire process from days to a few hours.

Glucose diffusion characteristics were not as accurate as the ones found in literature however the exact same conclusions were achieved and differences between calculated and literature results were minimal. By using the application to generate similar results for others OCAs and confronting those results to the ones presented in literature it will be possible to determine the true level of accuracy of the application, and if necessary improve it.

4.2 Future Implementations

Considering the field of optical clearing, the application functionality and purpose is considered to be up to date. However, there is a small exception. The forthcoming calculations regarding OCA viscosity in the diffusion process

that will be published soon by Oliveira and his research group which can be part of a following update to the application. In these calculations, these researchers have used the Stokes-Einstein equation (equation 10) to calculate OCA and water viscosity (η) value.

$$\eta = \frac{K_B}{6 \pi D r} \qquad (10)$$

In equation 10 the viscosity of an OCA or water can be calculated from the correspondent diffusion coefficient (D) value, if the Stokes radii (r) for these fluids are known. Boltzmann's constant, with the value of $k_B = 1.3807 * 10^{-23} J K^{-1}$ is also necessary in this calculation.

However, literature has not been updated yet relative to this calculation, and this will be implemented in the near future.

The cycle implementation to perform calculations (processing data) in all the experiments (concentrations) at the same time, is working better than was initial projected. On the other hand, the MATLAB tabs, which are the key for this concept, are not part of the components (such as text boxes, axes or tables), so it is more complicated to update each tab individually. At this point, every time a functionality is called a new tab-group is created, instead of refreshed. This happens because of the need for a dynamic number of tabs, making the entire tab-group access complicated. However a solution is already being projected, so the machine resources are not wasted.

In the meantime, other relevant doubts raised by users who tested the application can still be answered and improved. The time reduction to determine the diffusion times (using the application) have been achieved, and is now more dependent of the machine resources than the actual internal processing done by the implemented algorithms. This procedure created a small problem, quite common in fact and can be compared in our day-to-day

applications use. The user does not like to wait for the machine, and he is never sure if the machine is actually processing or not. For this reason the application log was added to the application, and the first tests were quite positive. However, the log must be more interactive, and be a part of the entire sequence (be present in both application windows, main and auxiliary).

The concept desired for the application was an unavoidable sequence of functionalities so the user always knows what to do. That means the objective is to always have that sequence implemented. The application can be used for other purposes if the functionalities are made with an higher level of freedom. This means that the application could be used in other concepts simply by changing the predetermined sequence. For example, load content into the application with the single purpose of using the maximum detection/selection, or normalizing procedures. The idea may seem rudimentary since the the application was developed only to determine the diffusion characteristics, but if the functionalities could be used in new sequences, selected through the interface, the application intent and objectives would increase and become more global. As an example, the application could be used for loading multiple matrices, where each matrix is a function with multiple lines, which could be represented in the tabs. A window containing these tabs can then be maximized for higher detail and edited without coding in MATLAB.

With this higher level of freedom, the functionality buttons can migrate to the menu on the top of the application. Allowing a new re-organization of the entire interface. This will not only allow the possibility to collapse the two windows (main and auxiliary) into one, but also show more information and options regarding the experiments that are being processed. As an example, discrimination of the two fluxes can be presented in the application interface (not being presented at the moment). Application log can include warnings for all functionalities since auxiliary window would be removed. New functionalities can be added since most buttons would migrate to the sub-menus.

As referred in section 3.6, another type of calculations/simulation will be considered for future versions of the application. As the diffusion properties for an OCA diffusion in a particular tissue are known, equations 1 to 6 can be used to simulate a treatment and determine the OCA concentration inside the tissue at any time of the treatment or the T_c time dependence can be generated. These simulated results can be very useful when a particular study is being prepared.

References

1. Oliveira L, Carvalho M I, Nogueira E and Tuchin V V 2015, Diffusion characteristics of ethylene glycol in skeletal muscle, *Journal of Biomedical Optics*, **20**(5), 051019-1-10

2. Oliveira L, Carvalho M I, Nogueira E and Tuchin V V 2013, The characteristic time of glucose diffusion measured for muscle tissue at optical clearing, *Laser Physics*, **23**(7), 075606-1-10

3. Genina E A, Bashkatov A N and Tuchin V V 2009, Glucose-induced optical clearing effects in tissues and blood, *in Handbook of Optical Sensing of Glucose in Biological Fluids and Tissues*, Tuchin V V (Ed) , pp 657-692, (CRC Press, Taylor & Francis Group, Boca Raton, Florida)

4. Tuchina D K, Shi R, Bashkatov A N, Genina E A, Zhu D, Luo Q, Tuchin V V 2015, Ex vivo optical measurements of glucose diffusion kinetics in native and diabetic mouse skin, *Journal of Biophotonics*, **8**(4), 332-346

5. Cheong W, Prahl S A and Welch A J 1990, A review of the optical properties of biological tissues, *IEEE Journal of Quantum Electronics*, **26**(12), 2166-2185

6. Wen X, Jacques S L, Tuchin V V, Zhu D 2012, Enhanced optical clearing of skin in vivo and optical coherence tomography in-depth imaging, *Journal of Biomedics Optics*, **17**(6), 066022-1-6

7. Genina E A, Bashkatov A N, Sinichkin Yu P, Yanina I Yu, Tuchin V V 2015, Optical clearing of biological tissues: prospects of application in medical diagnostics and phototherapy, *Journal of Biomedical Science and Engineering*, **1**(1), 22-58

8. Vargas G, Chan E K, Barton J K, Rylander H G and Welch A J 1999, Use of an agent to reduce scattering in skin, *Lasers in Surgery and Medicine*, **24**(2), 133-141

9. Zeng H, Wang J, Ye Q, Deng Z, Mei J, Zhou W, Zhang C and Tian J 2014, Study on the refractive index matching effect of ultrasound on optical clearing of bio-tissues based on the derivative total reflection method, *Biomedical optics Express*, **5**(10), 3482-3493

10. Vargas G, Barton J K, Welch A J 2008, Use of hyperosmotic chemical agent to improve the laser treatment of cutaneous vascular lesions, *Journal of Biomedical Optics*, **13**(2), 021114-1-8

11. Hirshburg J, Choi B, Nelson J S, Yeh A T 2006, Collagen solubility correlates with skin optical clearing, *Journal of Biomedics Optics*, **11**(4), 040501-1-3

12. Genina E A, Bashkatov A N and Tuchin V V 2010, Tissue optical immersion clearing, *Expert Review of Medical Devices*, **7**(6), 825-842

13. Barton J K 2011, Dynamic changes in Optical Properties, *in Handbook Optical-Thermal Response of Laser Irradiated Tissue*, (2nd Edition), Welch A J, Gemert M (Eds), pp 321-349, (Springer Netherlands)

14. Jiang J, Wanq R K 2004, Comparing the synergistic effects of oleic acid and dymethyl sulfoxide as vehicles for optical clearing of skin tissue in vitro, *Physics in Medicine and Biology*, **49**(23), 5283-5294

15. Bui A K, McClure R A, Chang J, Stoianovici C, Hirsburg J, Yeh A T, Choi B 2009, Revisiting optical clearing with dimethyl sulfoxide (DMSO), *Lasers in Surgery and Medicine*, **41**(2), 142-148

16. Mao Z, Zhu D, Hu Y, Wen X, Han Z 2008, Influence of alcohols on the optical clearing effect of skin in vitro, *Journal of Biomedical Optics*, **13**(2), 021104-1-6

17. Ribeiro A, Ortona O, Simões S, Santos C, Prazeres P, Valente A, Lobo V and Burrows H 2006 Binary Mutual Diffusion Coefficients of Aqueous Solutions of Sucrose, Lactose, Glucose, and Fructose in the Temperature Range from (285.15 to 328.15) K, *Journal of Chemical & Engineering Data*, **51**(5), 1836-1840

18. Ghosn M G, Carbajal E F, Befrui N A, Tuchin V V, Larin K V 2008, Differential permeability rate and percent clearing of glucose in different regions in rabbit sclera, *Journal of Biomedical Optics*, **13**(2), 021110

19. Bashkatov A N, Genina E A, Sinichkin Y P, Kochubey V I, Lakodina N A, Tuchin V V 2003, Glucose and mannitol diffusion in human dura matter, *Biophysical Journal*, **85**(5), 3310-3318

20. Bertram R and Pernarowski M 1998, Glucose diffusion in pancreatic islets of Langerhans, *Biophysical Journal*, **74**(4), 1722-1731

21. Zhong H Q, Guo Z Y, Wei H J, Zeng C C, Xiong H L, He Y H and Liu S H 2010, Quantification of glycerol diffusion in human normal and cancer breast tissues in vitro with optical coherence tomography, *Laser Physics Letters*, **7**(4), 315-320

22. Liu P, Huanq Y, Guo Z, Wanq J, Zhuanq Z, Liu S 2013, Discrimination of dimethyl sulfoxide diffusion coefficient in the process of optical clearing by confocal micro-Raman spectroscopy, *Journal of Biomedical Optics*, **18**(2), 020507

23. Ke M, Fujimoto S and Imai T 2013, SeeDB: a simple and morphology-preserving optical clearing agent for neuronal circuit reconstruction, *Nature Neuroscience*, **16**(8), 1154-1161

24. Ke M, Fujimoto S, Imai T 2014, Optical Clearing Using SeeDB, *Bio-Protocol*, **4**(3), 1-10

25. Grange R, Lanvin T, Hsieh C, Pu Y, Psaltis D 2011, Imaging with second-harmonic radiation probes in living tissue, *Biomedical Optics Express*, **2**(4), 2532-2539

26. Tuchin V V 2006, *Handbook: Optical Clearing of Tissues and Blood*, (Bellingham, WA: SPIE Press)

27. A. Kotyk and K. Janacek 1977, An Interdisciplinary Approach, in *Handbook Membrane Transport* (Plenum Press, New York)

28. Tuchin V V 2015, Light Scattering Methods and Intruments for Medical Diagnosis, *in Handbook Tissue Optics* (3rd edition, PM 254, Bellingham, WA: SPIE Press)

29. Kreft M, Lukšič M, Zorec T M, Prebil M and Zorec R 2013, Diffusion of D-glucose measured in the cytosol of a single astrocyte, *Cellular and Molecular Life Sciences*, **70**(8), 1483-1492

30. Tuchin V V, Maksimova I L, Zimnyakov D A, Kon I L, Mavlutov A H and Mishin A A 1997, Light propagation in tissues with controlled optical properties, *Journal of Biomedical Optics*, **2**(4), 401-417

31. Tuchin V V 1997, Light scattering study of tissues, *Physics Uspekhi*, **40**(5), 495-515

32. Bashkatov A N, Genina E A and Tuchin V V 2011, Optical properties of skin, subcutaneous, and muscle tissues: A review, *Journal of Innovative Optical* Health Sciences, **4**(1), 9-38

33. Jacques S L 2013, Optical properties of biological tissues: a review, *Physics in Medicine and Biology*, **58**(11), R37-R61

34. Mobley J, Vo-Dinh T and Tuchin V V, Vo-Dinh T (Ed) 2014, Optical Properties of Tissues , *in Handbook Biomedical Photonics*, 23-122, (Boca Raton FL: Taylor & Francis Group LLC CRC Press Inc.)

35. MATLAB 2015, Curve Fitting Toolbox, *in User's Guide*, 2.14-4.88, (MATHWORKS)

36. Reinoso R F, Telfer B A, Rowland M 1997, Tissue water content in rats measured by desiccation, *Journal of Pharmacological and Toxicological Methods*, **38**(2), 87-92

37. Gilard V, Martino R, Malet-Martino M, Riviere M, Gournay A, Navarro R 1998, Measurement of total water and bound water contents in human stratum corneum by in vitro proton nuclear magnetic resonance spectroscopy, *International Journal of Cosmetic Science*, **20**(2), 117-25