Solvatochromism as a new tool to distinguish structurally similar compounds

Pedro P. Madeira, Joana A. Loureiro, Mara G. Freire, João A.P. Coutinho

PII: S0167-7322(18)34600-2
DOI: https://doi.org/10.1016/j.molliq.2018.11.050
Reference: MOLLIQ 9960
To appear in: Journal of Molecular Liquids

Received date: 6 September 2018
Revised date: 12 November 2018
Accepted date: 13 November 2018

Please cite this article as: Pedro P. Madeira, Joana A. Loureiro, Mara G. Freire, João A.P. Coutinho, Solvatochromism as a new tool to distinguish structurally similar compounds. Molliq (2018), https://doi.org/10.1016/j.molliq.2018.11.050

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
Solvatochromism as a new tool to
distinguish structurally similar
compounds

Pedro P. Madeira*,a, Joana A. Loureirob, Mara G. Freirea and João A. P. Coutinhoa

*a CICECO - Aveiro Institute of Materials, Department of Chemistry, University of Aveiro, 3810-193 Aveiro, Portugal;
b LEPABE, Department of Chemical Engineering, Faculty of Engineering of the University of Porto, Porto 4500-465, Portugal.

*Corresponding author: Tel: +351 234401507; Fax: +351 234370084
E-mail address: p.madeira@ua.pt
Abstract

It is here reported a new concept based on solvatochromism to distinguish structurally similar compounds in aqueous solutions by the analysis of the stabilization of electronic excited states. The sensitivity of this approach to differentiate similar organic compounds, such as structural isomers or compound differing in the number of methylene groups, or proteins with conformational changes induced by being or not bound to cofactors, differing in two amino acids substitutions, or differing in their glycosylation profile, is demonstrated. The sensitivity of the proposed approach, based on the solvatochromic method, opens the path to its use as an auxiliary analytical tool in biomedical diagnosis/prognosis or in quality control of biologic-based drugs.

Keywords: Solvatochromic method • Structurally similar compounds • Biosimilars • Quality control • Biomedical diagnosis
Whilst recent advances in biotechnology are yielding many new products, the development of effective analytical and detection techniques lagged behind. Biosimilars represent an excellent example of why innovative detection techniques are needed. In 2010, worldwide sales of biologics approached the US$100 billion barrier\textsuperscript{1}, and by 2025 it is expected that more than 70\% of new drugs approval will be for biologics.\textsuperscript{2} As these drugs begin to come off patent, substantial opportunities exist to make copies (biosimilars) or generic versions of these drugs.\textsuperscript{3}

Every modern biomanufacturing process requires a complex set of analytical methods for product characterization, process validation, and quality control. For protein-based biopharmaceuticals, this is particularly relevant. Simple changes originated during the up-stream process, such as a single-amino-acid mutation or a post-translational modification, may result in a partial or complete loss of therapeutic activity.\textsuperscript{4-5} The higher-order structure of proteins is what gives each protein its three-dimensional shape and ultimately affects the way by which proteins act. The therapeutic properties of a macromolecule may significantly be impaired by changes in its 3D structure, and analytical methods to assess their high-order structures modifications are particularly advantageous. The monitoring of product variability to assess lot-to-lot consistency is also a purpose of modern analytical methods, mainly to assure that the chemical and 3D structure of the target product is maintained during manufacturing and purification processes.\textsuperscript{6} Any technique that provides information regarding the structural aspects, especially the structure and microheterogeneity of a product in a simple, rapid, and easy-to-understand format, is thus highly desirable.

Another area demanding for original detection approaches is the biomedical diagnosis/prognosis. Besides concentration, proteins in a biological fluid have multiple
characteristics that may change under different pathological conditions, including structural features, such as post-translational modifications, presence of single or multiple point mutations, truncations, etc. Methodologies that exploit these disease-induced changes in protein structure and interactions, thus creating molecular diagnostics with improved specificity and sensitivity, are highly demandable.

Current methods for biomolecules characterization include a host of sophisticated techniques, such as mass spectrometry (MS) and/or nuclear magnetic resonance (NMR) spectroscopy, which require highly experienced personnel and significant investment in instrumentation.

Herein, we report a new strategy based on solvatochromism to distinguish similarly related compounds in aqueous solutions by the analysis of the stabilization of electronic excited states. The term solvatochromism was introduced by Hantzsch to describe the changes observed in the UV-Vis absorption spectra of some specific dyes by variation of the solvent. These peak changes, to higher or lower wavelengths, reflect differential solvation of the ground and first excited state of the light-absorbing molecule, which correspond to shifts in the electronic transition energy. Several models have been developed to ascertain the excited state dipole moment of electronically excited molecules, which is quite useful, e.g., in designing nonlinear optical materials, in elucidation of the nature of the excited states and also in determining the course of a photochemical transformation.

The powerful influence of solvents on spectral absorptions is illustrated, for instance, by the intramolecular charge-transfer absorption band of the betaine dye 2,6-diphenyl-4-(2,4,6-triphenylpyridinium-1-yl)phenolate, whose spectra is shifted from \( \lambda_{\text{max}} = 810 \text{ nm} \) to \( \lambda_{\text{max}} = 453 \text{ nm} \) on going from diphenyl ether to water as solvent.
Such shift corresponds to a solvent-induced change in the excitation energy of ca. 28 kcal.mol\(^{-1}\).

Although solvatochromism has been mostly used to rationalize the solvent effects on reaction rates, solubility, distribution between two liquids, retention times in chromatography, rates of reactions, among others\(^{15-17}\), recently it has been applied to investigate aqueous solutions of different compounds as well as mixtures thereof.\(^{18-21}\) These were the first systematic solvatochromic studies on aqueous solutions covering a wide range of solutes of diverse chemical structures, such as inorganic salts, osmolytes, polymers, as well as proteins.\(^{18-21}\) These works have shown not only that solvatochromic methods could be applied to any aqueous solution, as long as the compound/mixture is water soluble, but also that solvatochromic studies in these solutions have potential applications beyond the physico-chemical characterization of these mixtures.

Here we investigate if solvatochromism has sufficient sensitivity towards solute structure, and if different experimental conditions can be manipulated to control this sensitivity. Solvatochromic dyes strongly absorb in the UV-Vis range, and in presence of distinct compounds will experience dissimilar intermolecular interactions displaying different solvatochromic behavior. The dyes chosen for the present work were 4-nitroanisole, \(\text{N,N-diethyl-4-nitroaniline}\), and 4-nitroaniline. Not only the electronic transitions of these dyes are responsive to different interactions in aqueous environment\(^{22-23}\), but also they have adequate water solubility, present intense bands transitions in experimentally accessible regions of the spectrum, and have positions of \(\lambda_{\text{max}}\) minimally influenced by band overlap.\(^{22}\)

In order to validate the experimental approach and the proposed concept, structural isomers were selected as the initial targets in proof-of-principle experiments.
We analyzed and compared the spectral data for the three above-mentioned dyes in aqueous solutions of simple alcohols (1-propanol vs. 2-propanol; 1-butanol vs. 2-butanol) and sugars (fructose vs. glucose), whose results are depicted in Figure 1.

![Figure 1](image1.png)

*Figure 1.* Potential of solvatochromic method to distinguish structural isomers in aqueous solutions. The energy transitions are given in kiloKaiser (kK) for the dyes/aqueous solutions.

In the aqueous solutions of the two pairs of isomeric alcohols being compared, the solvatochromic behaviors for the dyes 4-nitroanisole and N,N-diethyl-4-nitroaniline are identical, while the dyes shifts for 4-nitroaniline are slightly different. It should be mentioned that the former dye has been described to be more sensitive to solvent-acceptor to dye-donor hydrogen bond interactions. Our results are consistent with literature data,
in which isomeric alcohols are described to display different hydrogen bond acceptor basicities.\textsuperscript{24} Comparison of the dye’s solvatochromic behaviors in the carbohydrates solutions show that the peak shifts are identical for the dyes 4-nitroaniline and \textit{N,N}-diethyl-4-nitroaniline, while slightly different for 4-nitroanisole. These results are illustrated in Figure 1, showing that by the proper choice of the dye, the solvatochromic technique can be used to distinguish structural isomers in aqueous solutions, at concentrations below 600 milimolal (mmole.Kg\textsuperscript{-1}) (additional information can be found in the ESI – Tables S1-S3).

After appraising the potential of solvatochromism to distinguish between structural isomers, we then checked if the technique would have the required sensitivity to discriminate compounds differing in the number of methylene groups. For that end we analyzed and compared the spectral data for the same three above-mentioned dyes in aqueous solutions of simple alcohols (methanol, ethanol, 1-propanol, and 1-butanol), organic acids (formic, acetic, propionic and butyric acids), amines (n-propyl- and n-butylamine) and \(\alpha\)-amino acids (glycine and alanine). The respective results are shown in Figure 2. While no differences between the solvatochromic behavior of the dye \textit{N,N}-diethyl-4-nitroaniline are observed between aqueous solutions of glycine when compared to aqueous solutions of alanine, for all the other dyes-solutes combinations the technique successfully distinguishes homologous series of compounds in aqueous solution. By a proper choice of the dye, solvatochromism can be used to distinguish compounds, at concentrations below 600 mmole.Kg\textsuperscript{-1}, whose chemical structure differ only in a methylene group (additional information can be found in the ESI- Tables S1, S2, S4-S7).

We then evaluated if the sensitivity of the technique could be changed by the manipulation of the experimental conditions. Accordingly, we analyzed and compared
the spectral data of 4-nitroanisole, \( N,N \)-diethyl-4-nitroaniline, and 4-nitroaniline in aqueous solutions of fructose vs. glucose, and glycine vs. alanine, in aqueous solutions containing NaCl 0.15M, KCl 0.15M, NaSCN 0.15M, HCl (pH ~ 4), and additionally for glycine vs. alanine in NaOH (pH ~ 11) aqueous solutions.

Figure 2. Potential of solvatochromism to distinguish compounds differing in methylene groups in aqueous solutions. The energy transitions are given in kiloKaiser (kK) for the dyes/aqueous solutions.

It is known that in alkaline solutions, reducing sugars such as glucose or fructose exhibit instability.\(^{25}\) They display strong reducing intensity and become autooxidizable, becoming in a considerable extent interconvertible, polymerized and depolymerized. We evaluated the \( ^{13} \text{C} \) NMR spectra of fructose at alkaline solutions (Figure S1 in the ESI),
confirming the carbohydrate instability, and consequently no further studies were carried out under these conditions for the sugars solutions. Despite this, the results shown in Figure 3 demonstrate that the experimental conditions, such as presence of salt additives, can be judicious manipulated to increase the sensitivity of solvatochromism to distinguish similar related compounds, namely isomers such as glucose and fructose and compounds differing in the number of methylene groups such as glycine and alanine (additional information is given in the ESI - Tables S8 and S9).

Figure 3. Effect of the experimental variables on the sensitivity of the solvatochromic technique. Left figures reflect the behavior of the indicated dye in water of the indicated compounds, and right figures the solvatochromic behavior of the same dye in solutions containing, NaSCN 0.15M (sugars) or KCl 0.15M (amino acids).
It is relevant to notice that the spectrum of 4-nitroanisole is shifted from $\nu_{\text{max}} = 31.57 \text{kK}$ to $\nu = 31.55 \text{kK}$ when changing from an aqueous solution of fructose at 0.6 m to an aqueous solution of glucose 0.6 m. Such shift corresponds to an induced change in the excitation energy of the dye of ca. 51 cal.mol$^{-1}$. Comparatively, the excitation energy of the spectrum of the same dye is shifted from $\nu_{\text{max}} = 32.94 \text{kK}$ to $\nu_{\text{max}} = 32.89 \text{kK}$ on going from tert-butyl alcohol to ethanol as solvent, corresponding to a solvent-induced change of 143 cal.mol$^{-1}$. Moreover, the sensitivity of the technique can be increased by the addition of salts, solute concentration, or choice of dye. We analyzed the spectral data of 4-nitroanisole, $N,N$-diethyl-4-nitroaniline, and 4-nitroaniline, for the reasons previously mentioned, but the vast number of existing solvatochromic dyes yet to be tested, renders this technique as extremely promising as an auxiliary bioanalytical tool.

It is well recognized that proteins in a biological fluid present multiple characteristics that may change under pathological conditions, such as post-translational modifications, presence of single or multiple point mutations, truncations, oxidation, deamidation, phosphorylation, glycosylation, etc. It is thus possible to envisage the development of a methodology based on solvatochromism for diagnosis and/or prognosis based on the detection and monitoring of disease-relevant changes of proteins in biological fluids.

To test this hypothesis, we analyzed and compared the spectral data of the above-mentioned dyes in aqueous solutions of proteins. First we tested how the solvatochromic method would perform in aqueous solutions of proteins with conformational changes induced by being or not bound to cofactors. For instance, transferrins are iron-binding blood plasma glycoproteins that control the level of free iron (Fe) in biological fluids. When not bound to free iron ions, transferrin is known as "apotransferrin". We also tested
the technique in aqueous solutions of other proteins differing in two amino acids substitutions, namely with bovine β-lactoglobulin isoform A that differs from isoform B by two amino acid residues out of the 162 residues overall — isoform A has Asp-residue in position 64 and Val-residue in position 118, while isoform B has Gly-residue in position 64 and Ala-residue in position 118. Finally we tested the solvatochromic method as a bioanalytical tool in proteins differing in their glycosylation profile (Ribonucleases A and B). Bovine pancreatic ribonuclease exists in glycosylated (RNase B) and non-glycosylated (RNase A) forms. Ribonuclease B is a form of the enzyme RNase A that has an added glycoprotein with N-linked carbohydrates at Asn-34 linked to oligosaccharide containing 5–9 residues of mannose and 2 residues of N-acetylglucosamine per molecule.

The solvatochromic data obtained in aqueous solutions of the studied proteins are illustrated in Figure 4 (additional information can be found in the ESI, Tables S10-S12). In all the tested pairs of proteins, the solvatochromic method is able to discriminate proteins with different structures. It should be mentioned that no significant differences between the polypeptide structures in RNase A and B are detected by other techniques, such as X-ray analysis. The results for the transferrin’s are in agreement with the results obtained by partitioning in aqueous two-phase systems, whereas the results for the β-lactoglobulins A and B isoforms are in agreement with the results obtained by partitioning in aqueous two-phase systems, gradient chromatofocusing and electrophoretic, spectroscopic, and computational studies of the isoforms, all used as different approaches for identifying structurally similar proteins.
Figure 4. Illustration of the technique to distinguish proteins with conformational changes induced by being or not bound to cofactors, differing in two amino acids substitutions, or differing in their glycosylation profile.

The results here reported show that solvatochromism can be used as a new method able to distinguish compounds that differ only in a methylene group, structural isomers, as well as structurally similar proteins. The sensitivity of the method can be adjusted by judicious choice of the dyes, or of the experimental conditions, such as pH or use of salt additives. Finally, the proposed method is simple, of low cost, fast, presents high sensitivity, can be operated in high-throughput platforms, and requires minimal background training. The sensitivity of this approach, based on the analysis of the stabilization of electronic excited states, opens the path to its use as an auxiliary analytical tool in biomedical diagnosis/prognosis or in quality control of biologic-based drugs.
Experimental Section

Materials

Compounds

Glycine (lot #AO359169, purity >99%) and L-Alanine (lot #AO373124, purity ~99%) were purchased from Acros Organics. D-Fructose (lot # 142070200, purity > 98%) was purchased from Panreac AppliChem. D-Glucose (lot # 83H09561, purity > 99.5%) was purchased from Sigma Aldrich. Methanol (lot #1708961, HPLC grade) and ethanol (lot #1668935, analytical reagent grade) from Fisher Scientific, n-propanol (lot #0028/7, analytical reagent grade) from Lab-Scan, i-propanol (lot #1552099, HPLC gradient grade) from Fisher Scientific, n-butanol (lot #13L160501, analytical reagent grade) from VWR, i-butanol (lot #SZBE2460V, purity ≥ 99%) from Sigma-Aldrich, formic acid (lot #131030.1611, purity ≥ 98%) from Panreac, acetic acid (lot #21.0490508.500, purity ≥ 99%) from José M. Gomes dos Santos, propionic acid (lot #AO376752, purity ≥ 99%) from Acros Organics, butyric acid (lot #72740, purity ≥ 99%) from Riedel-de-Haën, propylamine (lot #10176635, purity ≥ 98%) and butylamine (lot #10186897, purity ≥ 99%) from Alfa Aesar were used without further purification. Human transferrin (lot #BCBS1576, purity ≥ 98%), Apotransferrin (lot # SLBS6313, purity ≥ 98%), β-
lactoglobulin A from bovine milk (lots # 120H800 and # SLBS3561V, purity ≥ 90%), β-lactoglobulin B from bovine milk (lot # SLBP8396V, purity ≥ 90%), Ribonuclease A from bovine pancreas (lot # SLBQ0318V, purity ≥ 60%) and Ribonucleases B from bovine pancreas (lot 017K7017, purity ≥ 60%) were purchased from Sigma Aldrich. All proteins were used without further purification. All solutions were prepared in water or aqueous salt solutions and used within 48h.

**Solvatochromic dyes**

p-Nitroanisole (1) (purity > 99%) was purchased from Sigma-Aldrich. The displacements for this dye are expected to characterize interactions with dye dipoles and induced dipoles.\(^{22}\) \(N,N\)-diethyl-4-nitroaniline (2) (purity > 99 %) was purchased from Fluorochem. The electronic transitions for dye (2) are expected to be more sensitive to solvent-donor to dye-acceptor hydrogen bond interactions.\(^{23}\) 4-nitroaniline (3) (purity > 99 %) was purchased from Sigma-Aldrich. The electronic transitions for dye (3) are expected to be more sensitive to solvent-acceptor to dye-donor hydrogen bond interactions.\(^{22}\)

**Other chemicals**

All salts and other chemicals used were of analytical-reagent grade and used without further purification. The water used was ultra-pure water, double distilled, passed
by a reverse osmosis system and further treated with a Mili-Q plus 185 water purification apparatus.

Methods

UV-Vis spectroscopy

The dyes 4-nitroanisole (1), N,N-diethyl-4-nitroaniline (2) and 4-nitroaniline (3) were prepared in water and used within 48h. Aqueous solutions (ca. 0.25 mg/mL) of each dye were prepared, and aliquots of 20-100 μL of each dye were added separately to a total volume of 1200 μL of a given solution. The samples were mixed thoroughly in a vortex mixer, except for the proteins solutions where gentle mixing was applied. At least 3 aliquots from each sample and 2 blanks were dispensed with a Multipette Xstream pipette (Eppendorf, Hamburg, Germany) to microplate wells. Following moderate shaking for 30 min in an incubating microplate shaker with temperature control (VWR, Pennsylvania, USA) for 30 min at 303K, the absorption spectra were acquired. To evaluate reproducibility, the maximum wavelength of each sample was evaluated in at least 3 separate aliquots, and in some instances at different days and using different lots. A BioTeck Synergy HT microplate reader with a bandwidth of 2.0 nm, data interval of 1 nm, high-resolution scan (~0.5 nm/s) and temperature control was used for the acquisition of the UV–Vis absorbance data. The absorption spectra of the probes were determined
over the spectral range from 210 to 550 nm in each solution. Pure solutions containing no dye (blank) were scanned first, at least in duplicate, to establish a baseline. The wavelength of maximum absorbance in each solution was determined as described previously\textsuperscript{33-34} using the PeakFit software package (Systat Software Inc., San Jose, CA, USA) and averaged. Standard deviation for the measured maximum absorption wavenumber was always ≤ 0.004 kK, and in most cases ≤ 0.002 kK for all dyes in all solutions examined.

\textsuperscript{13}C NMR

NMR data were obtained in ppm using a Bruker Avance 300 spectrometer (operating at 300.13 MHz for \textsuperscript{1}H and 75.47 MHz for \textsuperscript{13}C NMR). Solutions to be characterized, and a solution of tetramethylsilane (TMS) in pure deuterated water (99.9\% D) as internal standard, were used in NMR tubes adapted with coaxial inserts. A TMS/D\textsubscript{2}O solution was always used as the inner part of the concentric tubes, while each sample was used in the outer part of the NMR tube. Using this approach it is possible to guarantee that the TMS standard and D\textsubscript{2}O are not in direct contact with the sample, avoiding thus possible interferences or deviations in the \textsuperscript{13}C NMR chemical shifts. The Mnova software (Santiago de Compostela, Spain) was used for data processing. In Fig. S1 are represented two examples of NMR spectra.

Acknowledgements
This work was developed in the scope of the project CICECO-Aveiro Institute of Materials (Ref. FCT UID /CTM /50011/2013), financed by national funds through the FCT/MEC and when applicable co-financed by FEDER under the PT2020 Partnership Agreement. The authors acknowledge FCT for the post-doctoral fellowship SFRH/BPD/111113/2015 of P.P. Madeira. The authors acknowledge the financial support from the European Union Framework Programme for Research and Innovation HORIZON 2020, under the TEAMING Grant agreement No 739572. The Discoveries CTR. M.G. Freire acknowledges the European Research Council under the European Union’s Seventh Framework Programme (FP7/2007–2013)/ERC grant agreement no. 337753. This work was also supported through the project POCI-01-0145-FEDER-006939 funded by the European Regional Development Fund (ERDF), through COMPETE2020 and by national funds, through FCT - Fundação para a Ciência e a Tecnologia; and NORTE- 01- 0145- FEDER- 000005 – LEPABE-2 ECO-INNOVATION, supported by North Portugal Regional Operational Program (NORTE 2020), under the Portugal 2020 Partnership Agreement, through the European Regional Development Fund (ERDF).

References


28. Baek, W.-O.; Vijayalakshmi, M. A., Effect of chemical glycosylation of RNase A on the protein stability and surface histidines accessibility in immobilized metal ion


Highlights

- Solvatochromism is proposed as a tool to distinguish structurally similar compounds in aqueous solutions

- The method is able to discriminate closely related organic compounds in aqueous solutions

- The method is able to discriminate structurally similar proteins in aqueous solutions

- The proposed approach opens the path to its use as an auxiliary analytical tool.