BIOINFORMATICS APPLICATIONS NOTE Vol. 18 no. 1 2002
Pages 211–212

DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra

A. Lobley[1], L. Whitmore[1] and B. A. Wallace[1,2,*]

[1]Department of Crystallography, Birkbeck College, University of London, London WC1E 7HX, UK and [2]Centre for Protein and Membrane Structure and Dynamics, Daresbury Laboratory, Warrington WA4 4AD, UK

Received on May 16, 2001; revised on August 17, 2001; accepted on August 21, 2001

ABSTRACT

Summary: A user-friendly website for the analysis of protein secondary structures from Circular Dichroism (CD) and Synchrotron Radiation Circular Dichroism (SRCD) spectra has been created.

Availability: It is accessible via a server located at http://www.cryst.bbk.ac.uk/cdweb

Contact: Users may obtain information and apply for a UserID and password from cdweb@mail.cryst.bbk.ac.uk.

Supplementary information: This website forms part of the electronic resources and information service of the Centre for Protein and Membrane Structure and Dynamics (CPMSD) located at http://www.srs.dl.ac.uk/VUV/CD/cpmsd.html.

Circular Dichroism (CD) spectroscopy is an important method for examining protein secondary structure. A large number of algorithms and reference databases have been produced for the empirical analysis of percentages of alpha-helix, beta-sheet, beta-turns and other types of secondary structures from CD spectral data. Usage of, and facile comparisons between, the presently-available algorithms are difficult due to: input spectral formats that differ from each other and from the types of outputs of most commercial CD instruments, input formats requiring different scaling factors or expression in different types of units, the use of different reference databases, and the lack of comparability of different results due to the wide range of different statistical or goodness-of-fit parameters calculated by the different programs. The goal of the present work was to create a user-friendly website that would accept a variety of input and output formats and units, allow all the methods to use a range of common databases and permit facile comparisons between the results from different methods using graphics outputs and a consistent parameter to measure the comparability of experimental and calculated spectra in all cases (Lobley and Wallace, 2001).

DICHROWEB is a website established to facilitate such on-line analyses of protein secondary structures from CD spectra. It is available via a password-accessible server, which eliminates the problems of installation and software and hardware compatibility for casual users. It accepts as input a large number of standard ASCII output formats from a variety of commercial CD instruments (at the present time these include Aviv and JASCO instruments, with others to be included in later versions), output from the Daresbury Synchrotron Radiation Circular Dichroism (SRCD) instrument (Clarke *et al.*, 2000), a generic (spreadsheet-compatible) format, and both YY and BP output formats from the data processing program SUPER3 (Teeters and Wallace, 1986). It will accept input in a variety of the different units used to measure and report CD data (including mean residue ellipticity, molar ellipticity, delta epsilon, and theta). Data analyses can be done using a number of the most popular secondary structure determination algorithms, including SELCON3 (Sreerama and Woody, 1993; Sreerama *et al.*, 1999), CONTIN (Provencher and Glockner, 1981; Van Stokkum *et al.*, 1990), VARSLC (Compton and Johnson, 1986; Manavalan and Johnson, 1987), CDSSTR (Manavalan and Johnson, 1987; Sreerama and Woody, 2000), and K2D (Andrade *et al.*, 1993). It permits the use of a wide range of different reference databases (Sreerama and Woody, 2000; Sreerama *et al.*, 2000) with these algorithms. The reference databases encompass a number of different wavelength ranges, but the program limits the analyses to only those reference databases whose wavelength ranges are equal to, or narrower than, the extent of the wavelength range of the data to be analyzed. Hence, SRCD data sets which encompass a very large wavelength range (Wallace, 2000) can be suitably analyzed by all methods and databases. However, very limited wavelength range data sets are specifically

*To whom correspondence should be addressed (at Birbeck College).

excluded because they result in low accuracy results (Toumadje *et al.*, 1992). While the databases currently available are not suitable for the analyses of membrane proteins, work is underway to create a new membrane protein reference database from the spectra of membrane proteins whose crystal structures have been determined (Wallace *et al.*, unpublished results).

For comparisons of experimental and calculated spectra, DICHROWEB calculates a standard goodness-of-fit parameter, the Normalized Root Mean Square Deviation (NRMSD) (Mao *et al.*, 1982; Wallace and Teeters, 1987) which permits simple comparisons between all the different types of analyses and reference databases, and produces overlaid plots of experimental and back-calculated spectra. In addition, outputs of the calculated spectra are produced in several formats compatible with other standard plotting or spreadsheet programs. An on-line users' guide is provided and a number of precautionary warnings regarding interpretation of the results are noted. The website also includes information on the methods of analysis, definitions of the various units used in CD analyses, and links to other CD analysis and information sites. Thus it is anticipated that DICHROWEB will meet the needs of the CD users community in providing a simple, easily usable website that for the first time permits easy comparisons between results obtained using different methods of analyses and reference databases.

ACKNOWLEDGEMENTS

Supported by a project grant from the BBSRC. The Centre for Protein and Membrane Structure and Dynamics (CPMSD) is supported by a BBSRC Structural Biology Centres grant. We thank Drs Andy Orry, Bob Sarra, and David Clarke for help in testing this website.

REFERENCES

Andrade,M.A., Chacón,P., Merelo,J.J. and Morán,F. (1993) Evaluation of secondary structure of proteins from UV circular dichroism using an unsupervised learning neural network. *Protein Eng.*, **6**, 383–390.

Clarke,D.T., Bowler,M.A., Fell,B.D., Flaherty,J.V., Grant,A.F., Jones,G.R., Martin-Fernandez,M.L., Shaw,D.A., Todd,B., Wallace,B.A. and Towns-Andrews,E. (2000) A high aperature beamline for vacuum ultraviolet circular dichroism on the SRS. *Synchrotron Radiation News*, **13**, 21–27.

Compton,L.A. and Johnson,W.C. Jr. (1986) Analysis of protein circular dichroism spectra for secondary structure using a simple matrix multiplication. *Anal. Biochem.*, **155**, 155–167.

Lobley,A. and Wallace,B.A. (2001) DICHROWEB: a website for the analysis of protein secondary structure from circular dichroism spectra. *Biophys. J.*, **80**, 373a.

Manavalan,P. and Johnson,W.C. Jr. (1987) Variable selection method improves the prediction of protein secondary structure from circular dichroism spectra. *Anal. Biochem.*, **167**, 76–85.

Mao,D., Wachter,E. and Wallace,B.A. (1982) Folding of the H^+-ATPase proteolipid inphospholipid vesicles. *Biochemistry*, **21**, 4960–4968.

Provencher,S.W. and Glockner,J. (1981) Estimation of globular protein secondary structure from circular dichroism. *Biochemistry*, **20**, 33–37.

Sreerama,N. and Woody,R.W. (1993) A self-consistent method for the analysis of protein secondary structure from circular dichroism. *Anal. Biochem.*, **209**, 32–44.

Sreerama,N. and Woody,R.W. (2000) Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON and CDSSTR methods with an expanded reference set. *Anal. Biochem.*, **287**, 252–260.

Sreerama,N., Venyaminov,S.Y. and Woody,R.W. (1999) Estimation of the number of helical and strand segments in proteins using circular dichroism spectroscopy. *Protein Sci.*, **8**, 370–380.

Sreerama,N., Venyaminov,S.Y. and Woody,R.W. (2000) Estimation of protein secondary structure from CD spectra: inclusion of denatured proteins with native protein in the analysis. *Anal. Biochem.*, **287**, 243–251.

Teeters,C.L. and Wallace,B.A. (1986) SUPER3: a suite of programs for the processing and analysis of circular dichroism spectra. *Manual, version 2.1*.

Toumadje,A., Alcorn,S.W. and Johnson,W.C. Jr. (1992) Extending CD spectra of proteins to 168 nm improves the analysis for secondary structures. *Anal. Biochem.*, **200**, 321–331.

Van Stokkum,I.H.M., Spoelder,H.J.W., Bloemendal,M., Van Grondelle,R. and Groen,F.C.A. (1990) Estimation of protein secondary structure and error analysis from circular dichroism spectra. *Anal. Biochem.*, **191**, 110–118.

Wallace,B.A. (2000) Synchrotron radiation circular dichroism spectroscopy as a tool for investigating protein structures. *J. Synch. Rad.*, **7**, 289–295.

Wallace,B.A. and Teeters,C.L. (1987) Differential absorption flattening optical effects are significant in the circular dichroism spectra of large membrane fragments. *Biochemistry*, **26**, 65–70.