
CDPRO USER MANUAL

by E.Bucci

Note:

As reported by the authors, the CDPro software package consists of three of the popular programs for analysing the protein CD spectra for determining the secondary structure fractions: SELCON3, CDSSTR and CONTIN.

Supported OS:

The programs are written in FORTRAN and in a PC they work in MSDOS environment.

To install the programs:

- 1- Create a directory (e.g. c:\cdpro);
- 2- Copy or unzip all files in this directory.

To create an input file for the deconvolution:

- 1- Open the CD file you want to analyse with the Jasco program "Standard analysis";
- 2- Transform the CD spectrum in a molar ellipticity spectrum (menu Mathematics - Optical constant);
- 3- Divide the obtained spectrum for the number of protein residues (menu Mathematics – Arithmetics w. constant);
- 4- Execute the “dump” routine (menu Spectra - Dump) with a 1 nm sampling interval and a 185-260 nm spectral window;
- 5- Press the copy button on the results box;
- 6- Paste the data in an Excel worksheet;
- 7- Clear all the rows containing non-numerical characters and clear also the third column of data;
- 8- Save the obtained data file as a text file (not in the Unicode format!);
- 9- Close Excel and open the new data file by Word;
- 10- Change the tab characters into 4 spaces (use the Word search&change function and specify “^t” as text to search for and “ ” as text to insert);
- 11- Save the file as simple text in the CDPRO directory;
- 12- Run the CRDATA program to create the input file for the deconvolution;
- 13- The following lines should appears in the DOS shell:

```
Do you want to create a new INPUT file?
```

```
    If you want to CREATE a NEW INPUT file, TYPE 0
```

```
    If you want to APPEND the existing INPUT file TYPE 1
```

Press 0 and then Return;

-
- 14- The following line should appear:

Enter TITLE for your data - 40 characters

Input the name of the file containing your data (e.g. CD.txt);

- 15- The following lines should appear:

The number of lines to be skipped in CD file
Enter the number of CD values per nm
IF the DATA is .. 240nm,241nm, 242nm etc. enter 1
IF the DATA is .. 240nm, 240.5nm,241nm etc. enter 2
IF the DATA is .. 240nm, 240.2nm, 240.4nm etc. enter 5

Depending on your spectrum resolution, choose 1, 2 or 5 (number of lines per nm in the data file);

- 16- The following line should appear:

INPUT INITIAL wavelength:

If your spectral range is as specified in step 4, answer 260;

- 17- The following line should appear:

INPUT FINAL wavelength:

If your spectral range is as specified in step 4, answer 185;

- 18- The following lines should appear:

Is the CD data in Molar Ellipticity Units?
IF IT IS TYPE 1; IF NOT TYPE 0
IF YOU TYPE 1 DATA WILL BE CONVERTED TO DELTA(e)

Choose 1;

- 19- The following line should appear:

ASCII-file name (CD data)-MAX of 12 letters

Enter a file name for the input file to be created (e.g., MYSPEC.INP).

To deconvolute your spectrum:

- 1- First of all, you have to choose a proper reference set for the deconvolution. As reported by the authors, many reference sets of proteins from different sources are combined to create a large reference set of 48 CD spectra. Depending on your spectrum wavelength range the number of proteins in the reference set (IBasis) could be as large as 48. The reference sets are:
-

IBasis	Wavelength Range	Proteins	References	Secondary Structures
1	178-260 nm	29	(Johnson et al.)	6 (α R, α D, β R, β D, T, U)
2	178-260 nm	22	(Johnson)	6 (α , 3/10, β , T, P2, U)
3	185-240 nm	37	(Johnson et al., Keiderling et al., & Sreerama et al.)	6 (α R, α D, β R, β D, T, U)
4	190-240 nm	43	(Johnson et al., Keiderling et al., Yang et al., & Sreerama et al.)	6 (α R, α D, β R, β D, T, U)
5	178-260 nm	17	(Johnson et al.)	5 (α , β , T, P2, U)
6	185-240 nm	42	(Johnson et al., Keiderling et al., & Sreerama et al.)	6 (α R, α D, β R, β D, T, U)
7	190-240 nm	48	(Johnson et al., Keiderling et al., Yang et al., & Sreerama et al.)	6 (α R, α D, β R, β D, T, U)

- 2- Open the file “input” in the CdPro directory. Change the Basis_1 value in the “input” file in accordance with the IBasis set you prefer (from 1 to 7).
- 3- Run Selcon3;
- 4- You should find 4 new files in the CdPro directory, namely:
 1. Selcon3.out Summary of SELCON3 output
 2. CalcCD.out CD spectra reconstructed
 3. Basis CD and SS data of reference proteins
 4. Protss.OUT Detailed results appended to this file
- 5- Change the reference set in the “input” file and repeat step 3. Results from IBasis values 1, 3, 4, 6 and 7 should be comparable as the secondary structure fractions they estimate are identical (Kabsch & Sander, 1983, DSSP assignments as modified by Sreerama et al). IBasis = 2 (King & Johnson, 1999) and IBasis = 5 (DSSP, as modified by Sreerama & Woody, 1994) use a different assignment of secondary structure.
- 6- Run Continll and repeat steps 4-5;
- 7- Run CDSstr and repeat steps 4-5.
- 8- Compare the results obtained with different reference sets and the three different programs. The more they converged to one solution, the more you can be confident on the results (see also next section).

Deconvolution reliability:

Currently there is no “official” evaluation of the performances of the three software and the different reference protein databases included in the CDPro package documentation. However, in a recent paper (2000) the authors reported that “the performances of all three methods were comparable, in spite of the differences in the algorithms used in the three software packages. While CDSSTR performed the best with a smaller reference set and larger wavelength range, and CONTIN/LL performed the best with a larger reference set and smaller wavelength range, the performances for individual secondary structures were mixed. Analyzing protein CD spectra using all three methods should improve the reliability of predicted secondary structural fractions.” You can look to this paper also to have an idea of the error which can affect your prediction.

In my experience, however, the error on the predicted number of helices and beta strands is quite larger than that on the secondary structure predicted fractions, so that I would suggest to take with great care the numbers you get.
