Tutorial & Examples

Preamble

The N•ESPript interface consists of:

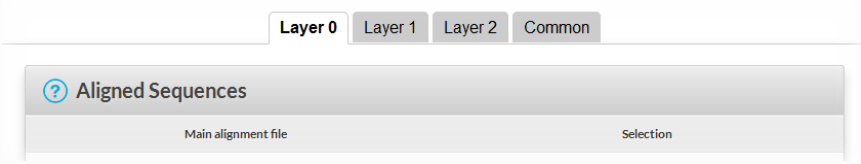
- A BUTTONS frame, fixed at the top of the page.
- A MAIN frame, which contains the user form.
- A POP-UP window containing the results of your **N•ESPript** job.

1 The MAIN frame:

- Fill up the form by uploading at least one multiple alignment file in the Aligned Sequences .
- PDB, CIF or DSSR files can be uploaded in the Secondary structure depiction section.
- The rest of the form allows you to change parameters related to the secondary structures and sequence similarity depiction, as well as the layout of the alignment output or the size and format of the resulting figures (PostScript, PDF, PNG, TIFF).
- All these options are explained in detail in the **User Guide** section or directly from the interface by clicking on the (?) icon.
- Tooltips are also available for form elements by moving the cursor over an (i) icon.

2 The BUTTONS frame:

- Only yellow buttons are active, with the exception of the TIME bar. Blue buttons are not clickable.
- When the main form is filled in, click on the SUBMIT button to let **N•ESPript** process your query.
- A result pop-up window will automatically appear within seconds. This results window can be (re-)opened at any time by clicking on the RESULTS button.
  - In order to access the results, you may need to **authorize your browser to display pop-up windows** from `nespript.ibcp.fr`. If necessary, consult your browser documentation or see our **F.A.Q. section** (second paragraph).
- The DOC button displays the full **User Guide** in a separate window.
- Click on ADV (ADVanced) or EXP (EXPert) to have access to more options. The default mode is BEG (BEGinner).
- In ADV mode, you can import another secondary structure file, change secondary elements labels and tinker with special characters.
- In EXP mode, you can also define your own colors, shift sequence numbering, etc.
- You can navigate between BEG , ADV and EXP modes without losing any information in your query.
- If you are in ADV or EXP mode, you can use the +1 button to build a layered **N•ESPript** figure. When you click on +1 , a new layer is created (up to 20 layers can be defined). The parameters set for Layer 0 (*i.e.* the first Layer) are copied to Layer 1 (and so on). Button -1 allows the user to suppress the last layer created. Layer Common contains parameters that are common to all layers.
- You can switch between the different layers by clicking on the tab bar (see below). See **example #3** to learn more about this option.



- Use the SAVE button to save your session to your computer, so that you can use the same parameters and files later.
- Use the LOAD button to load a previously saved session.
- Pay attention to the TIME bar. You must execute at least one command every 60 minutes otherwise your session will be closed.

- Before leaving, click on the EXIT button to permanently remove the processing and result files from the server.

3 The RESULTS pop-up window:

- Figures produced by **N•ESPript** appear in this multi-tab window with dedicated links. Left-click to visualise them, or right-click to retrieve them.
- Curious users can click on the Server logs | Input files or Tracing files tabs to access the corresponding reports and data.

4 EXAMPLE #1 - BEGinner Mode

This tutorial is based on the analysis of the crystal structure of yeast phenylalanine tRNA ($tRNA^{Phe}$), solved at 1.93 Å resolution (PDB entry **1EHZ** - Shi & Moore, 2000).
This structure is also used as an example in the manual of the program **X3DNA-DSSR** (Lu *et al.*, 2015), which is employed by **N•ESPript** to extract secondary structure elements from 3D nucleic acid structures using **Dot-Bracket Notation (DBN)**.

Session 1EHZ_beg.sav for beginners (resulting PDF or PNG files).

Sequences of known homologous structures to 1EHZ have been retrieved from the PDB using BlastN and aligned using ClustalW.
The **session file 1EHZ_beg.sav** allows the user to generate a figure from the alignment and to display secondary structure elements of 1EHZ in DBN. It was created as follows:

- Start **N•ESPript**.
- Save the **alignment file clustal-1EHZ.aln** on your disk and upload it in the first section Aligned Sequences of the form.

  Note: the multiple sequence alignment used (**clustal-1EHZ.aln**) contains DNA sequences, although they represent transfer RNAs (tRNAs). Activating the RNA button in the top section of the form converts thymines (**T**) into uracils (**U**), thereby transforming the DNA sequences into RNA sequences.
  Also note that lowercase characters in the alignment file **clustal-1EHZ.aln** indicate modified nucleotides. These characters are automatically converted to uppercase, unless you enable the Keep lowercase residues option in EXP mode.
- Check that the Global score in the Sequence similarities depiction parameters section is set to `0.7`. Click SUBMIT on the buttons frame and open the resulting image in the RESULTS window.

  A column framed in blue with a red background indicates strict identity, whereas a column framed in blue with a white background means that more than 70% of its residues are strictly conserved. Columns that do not meet these criteria are not framed.
  Note that 1EHZ corresponds to the fourth sequence in the alignment.
- Click the **PDB** logo displayed in the second section Secondary structure depiction (TOP secondary structures) of the form, and upload 1EHZ.
- SUBMIT and open the resulting image.

  Note that the secondary structure elements displayed above the sequence blocks are named 3EQ4_Y. This is normal: by default, they are aligned to the first sequence, here 3EQ4_Y. However, this is incorrect: for our case, the secondary structure elements should refer to 1EHZ.
- Scroll down to the Defining groups section of the form. Change the string `all to 4 all`
- SUBMIT and open the resulting image. 1EHZ is now displayed as the first sequence and secondary structure elements are correctly aligned.
- By default, the DBN coloring option is enabled in the Secondary structure depiction panel and the secondary structure elements in DBN are colored and labeled as follows:

  blue **( )** indicate regular base pairs, orange **( )** are wobble pairs; maroon **<>**, **{}**, **[ ]** are pseudo-knots; blue dots, labeled **η**, indicate hairpin loops; cyan dots, labeled **β**, indicate bulges; green dots, labeled **τ**, indicate internal loops; red dots, labeled **π**, indicate junctions; purple dots, labelled **α**, indicate non-loop single stranded segments.

  In addition, **–** indicate gaps and grey ★, modified bases if present in the PDB file.
- Click the Server logs tab in the RESULTS window to access to the list of information extracted by **X3DNA-DSSR** and displayed by **N•ESPript**. The full DSSR file can be downloaded from the Tracing files tab in the RESULTS window.
- Check the Alignments output layout section of the form. You will see a yellow tooltip with the message "*Suggested number of columns: 47 (for A4, portrait and a font size of 7)*". This indicates the Number of columns to use in order to produce a justified

figure.

- Change the Number of columns to 47, click SUBMIT and SAVE the session.

5 EXAMPLE #2 - ADVanced Mode

Session 1EHZ_adv.sav for advanced users (resulting PDF or PNG files).

This session uses the same ClustalW alignment **clustal-1EHZ.aln**, and PDB file, 1EHZ, as in **Example #1**.

- Start **N•ESPript** and click ADV in the buttons frame to access advanced options.
- Upload **clustal-1EHZ.aln** in the first section of the form and 1EHZ in the second section as in **Example #1**.
- Click SUBMIT and check the resulting image in the RESULTS window.

  Once again, **N•ESPript** incorrectly assumes that the uploaded PDB file corresponds to the first sequence in the alignment, whereas it actually refers to the fourth sequence (1EHZ). This time, we will correct the issue differently from **Example #1**, by using a new section called Special Commands and Characters, which is available in Advanced Mode ADV.
- Scroll down to the Special Commands and Characters section. You will see an image summarizing the various commands and colors available.
  X corresponds to secondary structure elements. Type the command `X R 1EHZ` to assign them to the sequence named `1EHZ` in the alignment, and to display the sequence name in red at the beginning of the displayed DBN.

  Remark: you can also type `X R 4` to assign the displayed DBN to the fourth sequence or type `X B 4` if you prefer the color blue.
  Click SUBMIT and check image. The adorned DBN now corresponds to 1EHZ.
- Go to the line following `X R 1EHZ` and type `Y R 1EHZ` to assign sequence numbering to the fourth sequence.
  Then, press *<return>* key and type `T R 1EHZ` to color the name of this sequence in red.
  Press *<return>* key once more and type `D R 35-37` to highlight the trinucleotide anticodon (residue 35-37 in 1EHZ numbering) with red downward pointing triangles.
- SUBMIT to verify.
- Enable the option Color by residues chemical properties in the section Special Commands and Characters and SUBMIT.

  Purine residues (**A**, **G**) are now written in white on a red background if strictly conserved. They are written in red on a grey background if >70% similarities and on a white background otherwise. The same scheme applies to pyrimidine residues (**C**, **T**, **U**), substituting red with blue. Remaining characters (here **N** for any base and **P** for pseudouridine) are in black.
- We will now take advantage of the fact that all names in the multiple sequence alignment correspond to PDB identifiers. Note that, for example, the sequence name `1I9V_A` means the A chain ID from the 1I9V PDB file.
  - Click Depict all known structures in the second section Secondary structure depiction and SUBMIT. **N•ESPript** will use **X3DNA-DSSR** to extract DBN for all aligned sequences based on their corresponding PDB file. Thus, DBN is displayed for most sequences, but you can also notice from the generated PDF file and the Server Logs tab in the RESULTS window that:
    - Displayed black dots correspond to low resolution structures for which only phosphorous traces (P-traces) are available.
    - **N•ESPript** currently do not support two characters chain IDs; Secondary structure elements are therefore not calculated for `pdb4v4r_aw`, `pdb8ccs_bb` and `pdb6gz3_bw`.
    - 2OW8 is an obsolete PDB entry that is no longer available from the current PDB database. Consequently, its DBN is not generated.
- Enable Depict all known structures except low-resolution ones and SUBMIT. DBN with black dots are removed.
- Finally, you can SAVE your session using the buttons frame for further modifications.

6 EXAMPLE #3 - EXPert Mode

Session 1EHZ_exp.sav for expert users (resulting PDF or PNG files).

Again, this session uses the same ClustalW alignment **clustal-1EHZ.aln**, and PDB file, 1EHZ, as in **Examples #1** and **#2**. We will now use the +1 button of the buttons frame to specifically display the secondary elements for the sequence `1I9V_A`. The following steps can be generated in either ADV or EXP mode and we will start from scratch.

- Open **N•ESPript** and choose either ADV or EXP mode.
- Upload **clustal-1EHZ.aln** in the first section of the form and 1EHZ in the second section as in Examples **#1** and **#2**.
- Type `X R 1EHZ` *<return key>* `Y R 1EHZ` in the Special commands and characters section as in **Example #2**.
- Enable Number sequences in the Aligned Sequences first section and click SUBMIT to check the result.
- We will now prepare the form for the +1 option by enabling Hide sequences in the first section and SUBMIT. Only 1EHZ

secondary structures elements are now displayed.

- Click on +1 in the buttons frame to create a new layer.

  The form now contains two independent layers named Layer 0 and Layer 1 as well as a shared layer (Common). Check that the alignment file, **clustal-1EHZ.aln**, has been copied from Layer 0 to Layer 1.

  IMPORTANT All subsequent commands must be entered in Layer 1.

- Untick Hide sequences to display sequence block in Layer 1.

- Scroll down and enter a Vertical shift of `-1` in the Alignments output layout section.

- SUBMIT to check the resulting image. You now have a gap between the secondary structure elements of 1EHZ (instructions in Layer 0) and the aligned sequences (instructions in Layer 1).

- Click the **PDB** logo displayed in the second section, Secondary structure depiction (TOP secondary structures) of Layer 1 and upload 1I9V. Then type `A` in Chain ID for selection.

- Type `X B 1I9V_A` *<return key>* `T B 1I9V_A` *<return key>* `T R 1EHZ` in the Special commands and characters section and SUBMIT to check.

- DBN labels of `1I9V_A` are unlikely visible. Enable Hide labels in the Secondary Structure section and SUBMIT.

  All information is now well displayed. We will finally add extra boxing with the `Q` special command to highlight specifically the anticodon of 1EHZ.

- This `Q` command, like `V` and `W`, uses column numbering instead of residue numbering. Click Ruler (preview for Q,V,W) in the Special commands and characters section of Layer 1 and SUBMIT to check column numbering.

- 1EHZ is the fourth sequence and its anticodon spans columns 39 to 41. In consequence, type `Q G 4 39-41` in the Special commands and characters section. Untick Ruler (preview for Q,V,W) in the Special commands and characters section to remove the ruler from the figure.

- Select the "Flashy mode" in the Color scheme section of the Common tab to add a yellow background on similar residues and modify the number of columns to 47 as suggested.

- SUBMIT, check final image and SAVE session.

7 EXAMPLE #4 - ADVanced Mode or EXPert Mode (recommended)

Session 1JJ2_9.sav for expert users (resulting PDF or PNG files).

You can use **N•ESPript** to directly read a **PDB** file, if you upload one instead of a multiple sequence alignment file in the first section of the form. Consequently, it can be used to quickly display the **DBN** of a given **PDB** structure.

For this final example, we will use chain ID 9 of the 1JJ2 PDB file (Klein *et al.*, 2001), which is also featured as an example in the **X3DNA-DSSR** manual (Lu *et al.*, 2015).

- Open **N•ESPript** in ADV mode.
- Upload 1JJ2 in the first section of the form and type `9` to select the chain ID (on the right side of the form).
- Upload 1JJ2 in the second section of the form, Secondary structure depiction, and type `9` to select the chain ID (just below).
- Click SUBMIT and check the resulting image in the RESULTS window.
- Enable the option Color by residues chemical properties in the Special commands and characters section.
- SUBMIT, check image and SAVE session.

8 References:

- Shi, H., and Moore, P.B. (2000) The crystal structure of yeast phenylalanine tRNA at 1.93 Å resolution: a classic structure revisited. *RNA*, **6**:1091-1105
- Lu, X.J., Bussemaker, H.J., and Olson, W.K. (2015) DSSR: an integrated software tool for dissecting the spatial structure of RNA. *Nucleic Acids Res.*, **43(21)**:e142
- Klein, D.J., Schmeing, T.M., Moore, P.B., and Steitz, T.A. (2001) The kink-turn: a new RNA secondary structure motif. *EMBO J.*, **20(15)**:4214-4221