

NATuG User Manual

Wolf S. Mermelstein and William B. Sherman

June 23, 2023

Contents

1	Getting NATuG Running	3
2	Quickstart	3
2.1	Running NATuG	3
2.2	Nucleic Acid Selection	4
2.3	Domain Adjustment	4
2.4	Junction Creation	4
2.5	Nicks, Linkages, and Sequences	4
3	Overall Layout	5
3.1	Hub	5
3.2	Panels	6
4	Config Panel	6
4.1	Graph Updating	7
4.2	Nucleic Acid Tab	7
4.2.1	Profile Manager	8
4.2.2	Setting Descriptions	8
4.3	Domains Tab	9
4.3.1	Settings Area	9
4.3.2	Angles Area	11
4.3.3	Counts Area	12
4.4	Sequencing	13
4.4.1	Bulk Sequence Operations	13
5	Side View Plot	13
5.1	Graphics	14
5.1.1	Axes	14
5.1.2	Points	15
5.1.3	Lines	15
5.2	Interaction	15
5.2.1	Panning	16
5.2.2	Zooming	16
5.2.3	Configuration	16
5.3	Modes	16
5.3.1	Informer Mode	16
5.3.2	Juncter Mode	17
5.3.3	Nicker Mode	18
5.3.4	Linker Mode	18
5.3.5	Highlighter Mode	19
5.4	Repetition	20
6	Top View Plot	20
6.1	Graphics	20
6.2	Interaction	21
7	Saving	21
8	Exporting	21
8.1	Export Settings	22
8.1.1	Side View Plot Settings	23
8.1.2	Top View Plot Settings	23

9	Strands	23
9.1	Selecting a Strand	23
9.2	Configuration	24
9.3	Sequence Selector	25
9.3.1	Manual Input Tab	25
9.3.2	Bulk Input Tab	26

1 Getting NATuG Running

The quick start guide is the fastest way to get NATuG running on any Mac, Windows, or Linux machine. These steps are by no means comprehensive, and are merely a guide to get the program running.

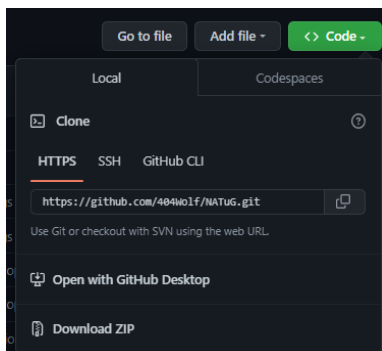


Figure 1: Github code download menu

1. Visit www.python.org/downloads, and then install the most recent version of Python for your operating system. NATuG has been confirmed to work on versions up to Python 3.11.1
2. Go to the github.com/404Wolf/NATuG, click the green “code” button, and then click “download ZIP.” See Figure ??.
3. Open your computer’s terminal/console/command prompt. Enter the following commands, in the following order.
 - (a) “`cd <filepath>`” to enter the directory of the project.
 - (b) “`python -m venv venv`” to create a virtual environment for the needed libraries to go into.
 - (c) “`venv/Scripts/activate`” if you are on windows, or “`source myenv/bin/activate`” if you are on Mac/Linux, to enter into the virtual environment. You will know that you have successfully entered the virtual environment if the current line in terminal begins with “(venv).”
 - (d) “`python -m pip install -r requirements.txt`” to automatically install all the needed libraries. They may take a while to download.
 - (e) “`python -m launcher`” to run the program. The first launch may take a bit while the code compiles.
4. You should now be in NATuG! When running the program in the future, simply CD into the folder (step 3a), enter the virtual environment (step 3c), and run the program (step 3e).

2 Quickstart

Constructing a DNA nanotube is a complex, multi-stage process, but NATuG streamlines the steps. Below is a quick list of things to do that should help you gain a basic level familiarity with NATuG.

Disclaimer: The below tasks are by no means comprehensive, and are meant to serve as a simple introduction to the program.

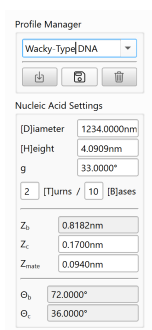
2.1 Running NATuG

First, go to Section 1 and follow the instructions to get NATuG running.

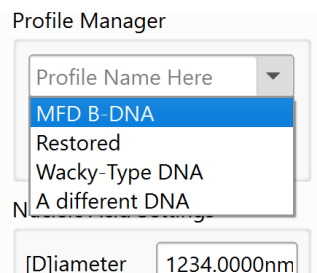
2.2 Nucleic Acid Selection

On the right side of the screen, in the config panel, you will find the Nucleic Acid settings tab. This is the area in which settings for the geometry of your DNA. While most of the time you will use NATuG’s default profiles for DNA settings (and, most of the time, the MFD B-Type DNA profile), now is a good time to experiment with the different inputs. As you change the different settings, also look at the bottom of the NATuG’s window—the Status Bar—when hovering over them. It should display a brief explanation of what the different settings actually are. Observe the Side View Plot and Top View Plot change as the settings update.

Figure 2: Experimenting with Nucleic Acid settings



(a) Tinkering With Nucleic Acid Settings



(b) Loading different Nucleic Acid profiles

2.3 Domain Adjustment

In the config panel on the right side of the screen, switch to the Domains tab. The panel will pop out because of the required width of the tab. The panel may appear intimidating at first, but you should start by using the small arrow buttons next to the "m" of each domain. As you do so, watch the Top View Plot update in real time. By increasing/decreasing the "m" values, the interior angles of domains will change, which will in turn change the shape of the tube. You can type numbers directly into the boxes, but without correcting for changes with other domains’ angles, the tube will open up. Notice how when the tube is closed the M/R box in the symmetry section of the tab lights up green.

In the settings area above the domains table, click the load button (the button that has an arrow pointing downwards into a box). Choose a different set of domains, for instance, "nested.csv." As you play around with this section of NATuG, you can use the save button to save your own designs/your adjustments to the default designs.

2.4 Junction Creation

In the Side View Plot—the main area of NATuG—locate two white triangles (which represent the middle of two nucleosides, and are called "NEMids") that are overlapping, and click on them. This will transform two helices of two different domains into two helices that traverse both of the domains. Continuing clicking on overlapping white triangles to weave together strands. See Figure ?? for a visual example.

2.5 Nicks, Linkages, and Sequences

Here, you can begin to experiment with more advanced features of NATuG. You just experimented with creating junctions, but the Side View Plot mode that creates junctions ("juncter" mode) is only one of many modes (though it’s likely the mode you’ll be using the most). To try out other modes, go to the top of the window, where you see a list of different buttons ("Informer," "Juncter," "Nicker," etc.), and click on different modes. Then left click on points within the main plot. Below is a list of (very) simplified descriptions of the different modes and how to get started using them.

Figure 3: Changing up domain settings

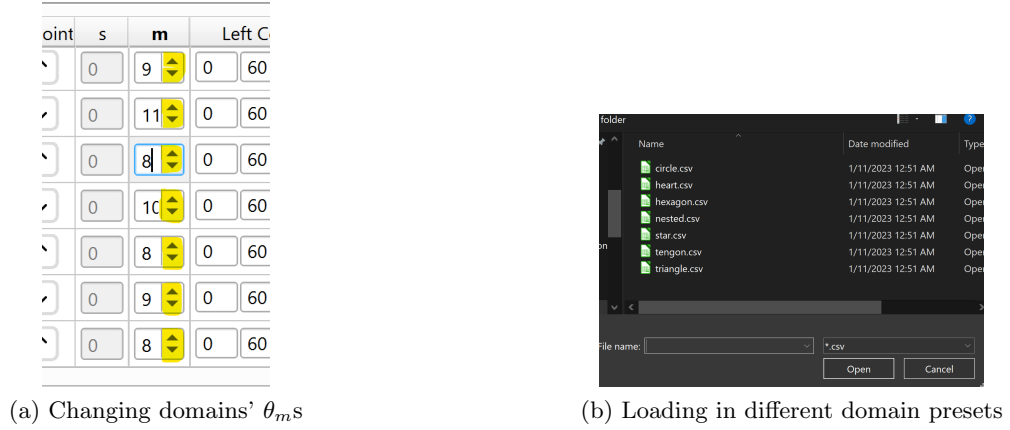
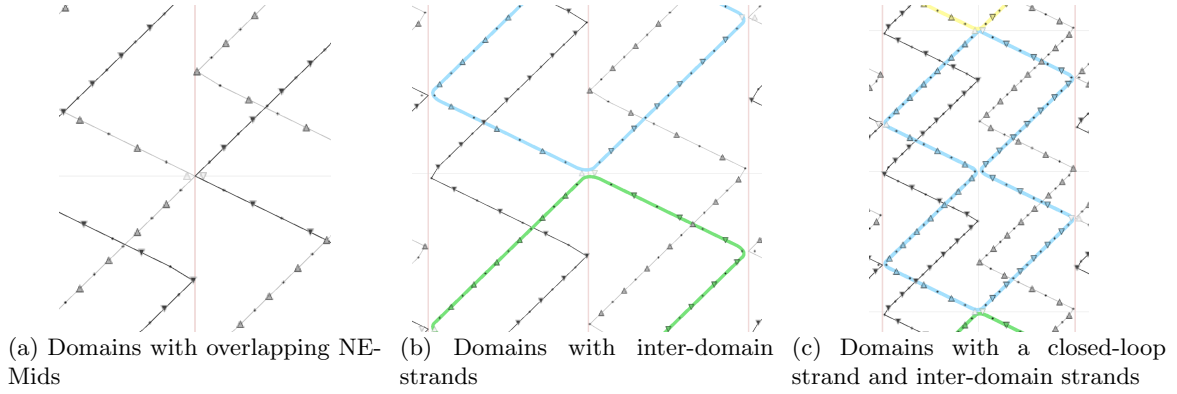


Figure 4: Creating cross-strand junctions



1. **Informer Mode:** Obtain information about points. To use, click on any point.
2. **Juncter Mode:** Redirect strands across helical domains. To use, click on any two overlapping white triangles.
3. **Nicker Mode:** Split strands in half. To use, click on any point.
4. **Linker Mode:** Link the ends of two strands together to make one bigger strand. To use, click on one point that is at the end of a strand, and then another point that is at the end of a strand. Make sure that the arrows do not point towards each other.
5. **Highlighter Mode:** Highlight points. To use, click on any point.

3 Overall Layout

NATuG's interface consists of a main hub area, surrounded by two panels. There is a status bar at the bottom of NATuG, which provides helpful descriptions of what various buttons and input boxes do as you hover over them, and a file bar at the top of NATuG, which provides access to various cross-program functions.

3.1 Hub

NATuG's main area, the "hub," is another name for the area of the program where the Side View plot is located. Located directly above the Hub is the toolbar, where you can select the current mode of NATuG.

Figure 5: Advanced NATuG features

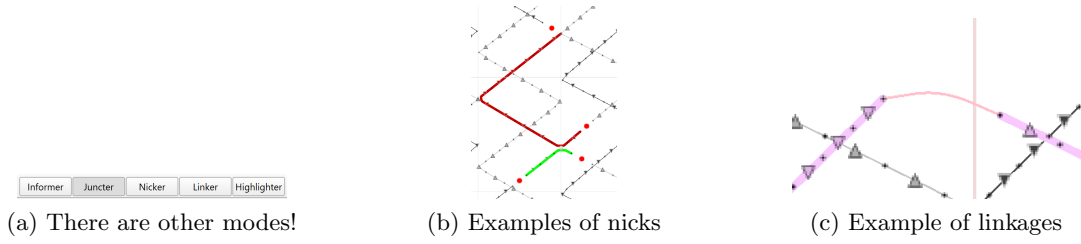
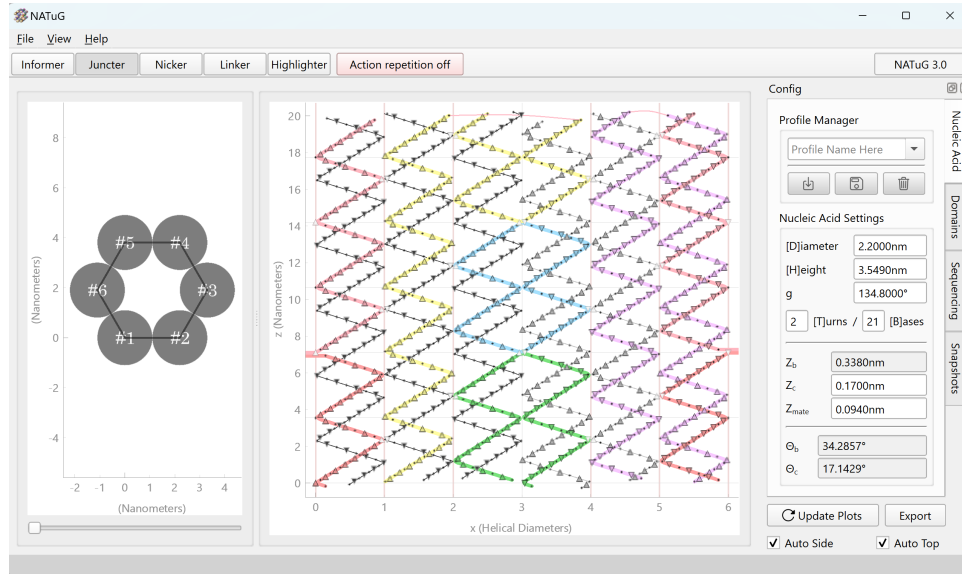


Figure 6: The overall layout of NATuG



Each mode changes what a left click does in the Side View plot.

3.2 Panels

NATuG currently has one Panel: the “config” panel; however, in the future additional Panels may be added. Some of the important properties of Panel-type widgets are listed below.

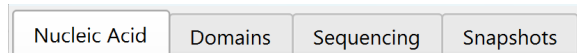
1. Panels are all undockable. By clicking twice on the top area of a panel, or by dragging the panel away from the main window, or by clicking the button, one can make a panel become its own window. To redock, click on the top of the panel twice, or drag the panel back into the main window.
2. Panels are hideable. By clicking the “x” button in the top corner of a panel, or by going to the file menu, “view,” and then the name of the panel you are trying to hide, you can hide the panel. To unhide a panel, go to the file menu, “view,” and then the name of the panel you are trying to unhide.

4 Config Panel

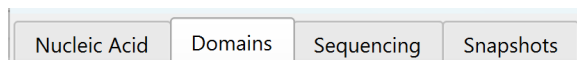
The Config Panel contains most of the user input areas of NATuG. Within the panel are three primary Subpanels: “**Nucleic Acid**,” “**Domains**,” and “**Sequencing**,” and “**Snapshots**,” outlined in greater depth on the following pages.

To navigate between Subpanels of the config panel, simply click on one of the tabs on the right. Below are brief descriptions of the various tabs; however, each tab has a dedicated page as well.

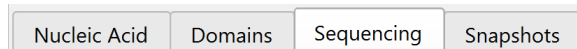
Warning: changing settings within the Config Panel will reset all current junctions, nicks, links, and highlights. Attempting to change settings when there exist any junctions, links, or highlights will result in a popup dialog that offers to save the current state of the Side View Plot and then refresh the plot, refresh the plot without saving, or cancel the attempted refresh. Thus, you generally want to configure the geometry of the structure before you interact with the strands.



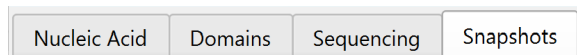
The Nucleic Acid Tab is where one can customize the geometrical settings of the nucleic acid that they are creating a nanotube for. They default to settings for a slight variation on B type DNA, and generally does not need to be changed.



The Domains Tab is where settings for the interior angles, strand switches, and the number of NEMids to generate are inputted. This section allows the user to actually define the shape of the nanotube.



The Sequencing Tab is where the user can apply bulk sequence actions to all the strands, and export sequences to a spreadsheet for synthesis. This is generally one of the later steps in the nanotube design process.



The Snapshots Tab is where version history can be found. It lets you rollback to previous program states with ease, and create “snapshots” of states at any time.

4.1 Graph Updating

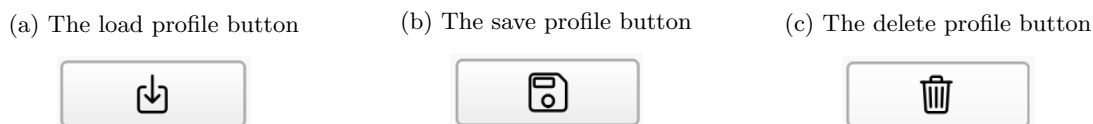
By default, as changes are made within the various tabs of the Config Panel the Side View Plot and Top View Plot automatically update. To disable automatic updating for either plot, click either of the “Auto Side” or “Auto Top” buttons on the bottom of the Config panel. When automatic updating is disabled, the “Update Graphs” button must be clicked to refresh the plots. Manual updating can be useful for larger structures that take a long time to load. Note that these automatic update buttons **only** apply to updates made within the Config Panel, so any changes made via strand interaction (i.e. such as creating junctions or linkages) will cause the plot to refresh to reflect the change made through the interaction, but may not be indicative of the current state of the Config Panel if auto updating is disabled and update plots has not been clicked.

The current tab of the Config Panel determines what is currently plotted in the Side View Plot. When the current tab is either “Nucleic Acid” or “Domains” NEMids are plotted, and the mode of the Side View Plot is unrestricted. However, when “Sequencing” is active, nucleosides are plotted. This difference is easy to notice, since the difference in the plots is that everything is slightly shifted vertically. When nucleosides are currently plotted (and, by extension, the “Sequencing” Subpanel is visible), only the “highlighter” and “informer” modes are enabled within the Side View Plot’s toolbar.

4.2 Nucleic Acid Tab

The Nucleic Acid Tab is the area in which all geometrical settings for the nucleic acid being used can be entered. All of NATuG’s computations utilize these constants, so only make changes if you know what you are doing. This tab contains a profile manager for easily loading in sets of settings, and has various fields

Figure 7: Various Profile Actions



that warrant extended descriptions. When using NATuG, recall that hovering over fields will show additional information in the Status Bar at the bottom of the window—this feature is particularly useful for the Nucleic Acid Tab because hovering over setting fields shows descriptions in the status bar of what they represent.

4.2.1 Profile Manager

The Profile Manager lets you easily save and load various sets of nucleic acid settings. Using NATuG’s profile defaults will ensure that all the settings are correct for the type of DNA you are working with, although you can also create your own profiles for future use.

By default, NATuG includes common profiles, such as one for B type DNA. Profiles are saved internally, and can easily be loaded/retrieved, and .json files containing the profile data can be found within the saves/nucleic_acid folder. These files are loaded when NATuG is booted, and can be modified directly. Additionally, when NATuG is closed, the last used settings are saved, and are automatically retrieved upon the next launch. The default profiles that ship with NATuG cannot be removed.

The Profile Manager allows you to load profiles, save profiles, and delete profiles. To do so, enter the name of the profile you would like to save/load/delete into the “Profile Name Here” box, and click the respective button. Below are extended descriptions of what the three actions do. When a Profile Manager box is disabled/gray, consult the status bar for information as to why (this can happen for various reasons, like trying to delete a default profile). For irreversible actions, such as deleting a profile, warnings will be shown before the action is run that will give you a chance to not do the action. Below are extended descriptions of the various Profile Manager actions.

Load Profile Load the profile with the currently chosen profile name. This requires a profile to already exist with the name chosen, and for the current settings to not be the settings of that profile. (Icon ??)

Save Profile Save the profile with the currently chosen profile name, or, if there is already a profile with that name, overwrite the existing profile. The profile will be saved to a .json file within saves/nucleic_acid after the program is closed. (Icon ??)

Delete Profile Delete the profile with the currently chosen profile name. This action is irreversible. Default profiles cannot be deleted. (Icon ??)

4.2.2 Setting Descriptions

Below lies a list of all of the various Nucleic Acid settings that can be adjusted. Certain settings (the ones that are have an ”(a)” in the table) are automatically determined based on other settings and cannot be user set.

Table 1: Nucleic Acid Field Descriptions

Input	Name	Data Type	Description
D	Diameter	number	The diameter of a given domain in nanometers
H	Height	number	The height of one turn about the helical axes

g	Nucleoside-Mate Angle	number between 0 and 360	The angle about the helical axis between a nucleoside and its Watson-Crick mate
T/B	Helical Turns per Bases	integers	There are T turns every B bases
Z_b (a)	Base Height	number	The height between two NEMids on a given helix
Z_c	Characteristic Angle	number	The height a helix climbs as it rotates through the characteristic angle
Z_{mate}	Nucleoside-Mate Height	number	Vertical distance between a NEMid and its mate on the other helix.
θ_c (a)	Characteristic Angle	number between 0 and 360	The smallest angle about the helical axis possible between two NEMids on the same helix.
θ_b (a)	Interbase Angle	number between 0 and 360	The angle that about the helical axis between two NEMids

4.3 Domains Tab

Figure 8: Domains Config Table

(a) Angles Tab of Domains Config Table

(b) Counts Tab of Domains Config Table

The Domains Tab is the area in which all the settings for each helical domain can be set. The angles between each domain are what determine what the ultimate shape of the nanostructure will be. The “counts” are what determine how many nucleosides each strand will be granted. Amongst many other design tools, NATuG allows for designing symmetrical structures, along with the ability to save and load designs. The tab consists of three components: the Settings Area (??), Angles Area (??), and Counts Area (??), all explained in greater detail below.

The Angles Area (??) and Counts Area (??) are placed within a tabbed area on the bottom part of the Domains Tab, such as not to overwhelm you with too many settings in the same table. Each row corresponds to the same domain, even though columns are bifurcated into two tabs.

4.3.1 Settings Area

The settings area’s primary purpose is to allow you to set the number of domains for your structure, however it also includes some critical parameters apropos symmetry.

Figure 9: Domains Subpanel Components

Settings

Domains (C)
Subunits (R)
Total (CR)

x =

M (Σm)
M/R
Target M/R

☒ Auto Antiparallel

Angles		Counts			
	L-Joint	R-Joint	s	m	θ_i
#1	↑	↑	0	7	120.00°
#2	↓	↓	0	7	120.00°
#3	↑	↑	0	7	120.00°
#4	↓	↓	0	7	120.00°
#5	↑	↑	0	7	120.00°
#6	↓	↓	0	7	120.00°

Angles		Counts					
		Up Helix Counts			Down Helix Counts		
#1		0	60	0	0	60	0
#2		0	60	0	0	60	0
#3		0	60	0	0	60	0
#4		0	60	0	0	60	0
#5		0	60	0	0	60	0
#6		0	60	0	0	60	0

(a) Settings area
(b) Angles area
(c) Counts area

Symmetry is particularly useful for designing closed structures. What it means, simply put, is that the angle and count settings of a cluster, or “subunit,” of domains is repeated a certain number of times. Because of this, the Subunit Table (4.3.2 & 4.3.3) is **not** a place to enter settings for every single domain, but rather contains the domains of a single subunit; single subunit structures are a special case in which the Subunit Table does showcase the settings of all domains. Within this settings area of the Domains Tab you can change the number of copies of the domains listed in Subunit Table, along with more advanced settings, listed below.

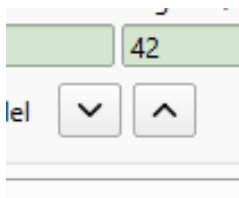
Once you have made changes within this area, click the check icon, the top button of the group of three buttons on the right, to confirm your choices for the number of domains. If you choose to reduce the number of domains per subunit, a warning will be displayed informing you that the subunit table is to be truncated to allow for the reduction.

Table 2: Settings Area Fields

Name	Datatype	Description
Domains	integer	The number of domains per subunit. This dictates how many domains show up in the Subunit Table. This is also known as the domain count, or N
Subunits	integer	How many times to copy the current subunit. This is also known as the symmetry factor, or R
Total	integer	The total number of domains within all subunits. This is equivalent to $R \times C$.
M	integer	The sum of all the ms for all the domains within all the subunits.
$\frac{M}{R}$	integer	The sum of all the ms for a given subunit. This is equivalent to, as the name implies, $\frac{M}{R}$.
Target $\frac{M}{R}$	number	This is equivalent to $B \times \frac{N-2}{2R}$, where B is a parameter of the specific nucleic acid (which can be adjusted within the Nucleic Acid Tab, discussed in more detail in section 4.2), C is number of domains, and R is the symmetry factor. If this is not an integer then that means that, by virtue of the chosen ms , forming a closed tube is impossible.

Antiparallel	boolean	Whether the connection between the right joint of the last domain to the left joint of the first domain in the next subunit should be forced to be antiparallel. If this is enabled and this condition is already met, no action is performed. If there is an odd number of subunits it is not possible to alternate parallelity of each subunit and achieve an antiparallel relationship between the right joint of the last domain of each subunit to the next subunit's first domain's left join. If this is the case an error will be raised, and alternation will occur notwithstanding.
--------------	---------	---

Figure 10: Rotate Domains Buttons



An additional tool worth noting here is the **rotate domains buttons**, shown in figure ???. These buttons allow you to shift the settings of each domain in the template subunit upwards one, where the first domain's settings take the place of the last domain's settings, or the other way around. These buttons are particularly useful when dealing with non-symmetrical structures.

4.3.2 Angles Area

The Angles Area of the Domains Subpanel is where you can enter information about the angle of helical domains in relation to other helical domains, along with controlling how the strands align with one another. Each row represents one of the helical domains within the **first** subunit, and the settings here are copied across all subunits.

Table 3: Settings Area Fields

Column	Datatype	Description
L-Joint	<i>UP</i> or <i>DOWN</i>	One of the two helices in the double helix is to be lined up to one of the two helices of the previous double helix. This determines whether to align this double helix's up or down helix to the previous domain's double helix.
R-Joint	<i>UP</i> or <i>DOWN</i>	Whether to align this double helix's up or down helix to the L-joint helix of the next domain's double helix.

s	integer	<p>θ-switch multiple. This is automatically determined based off of the L-joint and R-joint. $s \in \{-1, 0, 1\}$, and is determined as follows:</p> <ul style="list-style-type: none"> • $s = -1$ if the L-joint is <i>UP</i> and the R-joint is <i>DOWN</i>. • $s = 0$ if the L-joint's direction matches that of the R-joint. In this case there is no strand switch. • $s = 1$ if the L-joint is <i>DOWN</i> and the R-joint is <i>UP</i>. <p>It's easy to remember because if you're going <i>UP</i> to <i>DOWN</i> you're doing a negative strand switch, and the same logic applies the other way around.</p>
m	integer	<p>The number of characteristic angles between the domain's center and the next domain's center. This is a multiple of θ_c, (which can be set in the Nucleic Acid Tab, discussed in section 4.2). While you can manually enter an integer into this box, if you increase the angle of a single domain without compensating elsewhere, the tube will no longer be closed. Clicking the up/down-tick buttons will automatically adjust the interior angle of the surrounding domains, so that this one increases by 2, and the next and previous ones decrease by 1, and vice versa, to maintain the value of M.</p>
θ_i	number	<p>The actual interior angle between this and the next domain. Equivalent to $\theta_m + \theta_s$, which is also equivalent to $\theta_c * m + \theta_s * s$</p>

4.3.3 Counts Area

The counts area lets you set the number of nucleosides of each strand. All strands are required to begin and end with a nucleoside, with NEMids interspaced between every nucleoside. Because of this, the number of NEMids in each strand is equal to the number of nucleosides in the strand, which can be set in this area of the Domains Tab minus one.

Like the angles area (4.3.2), each row represents one of the helical domains within the **first** subunit, and the settings here are copied across all subunits. The number of nucleosides for the up helix of a given domain in the template subunit can be set by entering values into the respective row of the "Up Helix Counts" column. Likewise, you can set the number of nucleosides for the down helix by altering the values of the respective row within the "Down Helix Counts" column.

The area wherein you can enter the number of nucleosides you'd like for a given strand takes the form of three integers. They represent the following:

- **Left:** After generating the initial number of nucleosides (set in the middle box), this is how many additional nucleosides to generate and prepend to the bottom of the strand. This can be negative to trim NEMids.
- **Middle:** The initial number of nucleosides to generate for the given helix.
- **Right:** After generating the initial number of nucleosides (set in the middle box), this is how many additional nucleosides to generate and append to the top of the strand. This can be negative to trim NEMids.

Note that adjusting the left and right boxes can be particularly useful for creating sticky ends, since, when the number of additional nucleosides to generate atop a helix does not match that of its complementary helix, there then exist nucleosides that do not have mates (and could theoretically be given bases that correspond with some external viral strand).

4.4 Sequencing

Figure 11: Sequencing Tab

This tab allows for bulk sequence operations, and for exporting sequences to a spreadsheet for synthesis. Because sequencing is such a critical part of NATuG, it has a section that goes into greater depth about how it works. It is also important to note that this is not the only place where sequences can be set. For more information about sequencing, consult section 9.3.

4.4.1 Bulk Sequence Operations

Bulk sequence operations let you set the sequence of many nucleosides at once. In order to run a bulk sequence operation, select a "Scope," then an "Operation," and then click "Run."

Scopes The scope is which nucleosides the operation should run on.

- **All bases:** All the nucleosides of all the strands.
- **Unset bases:** All the nucleosides of all the strands that do not currently have a base set.

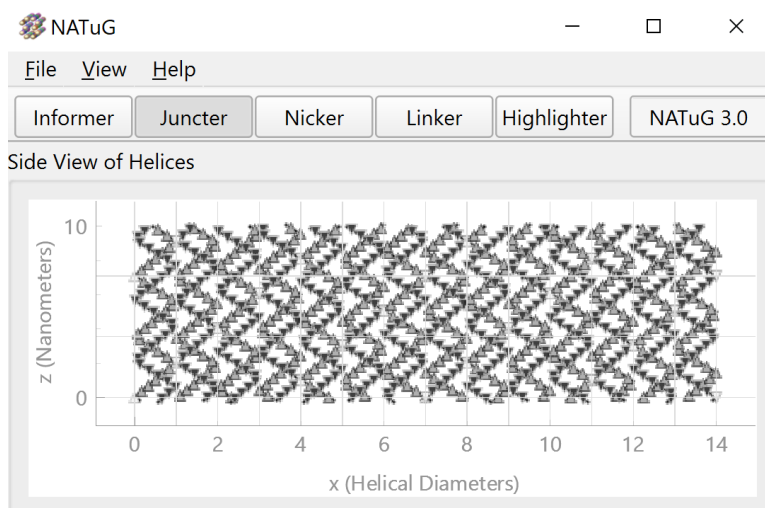
Operations The operation is what to do to the nucleosides within the scope.

- **Randomize:** Select random bases.
- **Clear:** Unset the bases.

5 Side View Plot

The Side View Plot is the heart of NATuG. It allows you to directly interact with strands that have been created as a result of user inputs set elsewhere. The plot can be broken into a few distinct parts that function together to allow for a high degree of interactivity.

Figure 12: The Side View Plot & toolbar



5.1 Graphics

Before diving into how to actually interact with the Side View Plot, it is important to understand what the contents of the plot actually represent. The Side View Plot is not, as its name suggests, a direct side view of the double helices. Rather, it is more of an “unrolled” plot, since it is distorted in such a way as to neatly lay out the double helices next to one another. It is a type of projection.

Importantly, the plot consists of various artifacts, which represent physical nucleosides, or abstract areas between physical nucleosides, NEMids. Below lies further descriptions of what these artifacts appear as.

5.1.1 Axes

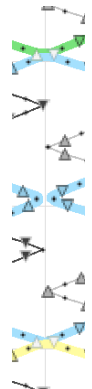
Axis units The x and y axes have labels that indicate their dimension, but the gridlines represent more intuitive concepts.

- The x -axis labels represent helical diameters, and thus also define sections between domains. All points between $x = 0$ and $x = 1$ are in domain#1, those between $x = 1$ and $x = 2$ are in domain#2, and so forth.
- The y -axis labels represent the actual vertical height, and the gridlines represent each complete helical twist.

(a) Unstable joint example (0 junctions)



(b) Stable joint example (2 junctions)

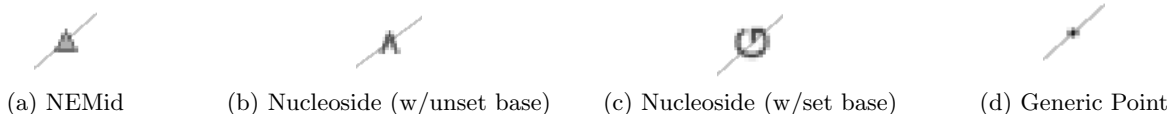


Gridlines The **horizontal** gridlines lie beginning at $z = 0$ and repeat upwards up until the top of the highest point (nucleoside/NEMid). The gridlines themselves are all gray and identical in style.

The **vertical** gridlines lie beginning at $x = 0$ and repeat up until the right of the last helical domain. For each helical joint, including the last one, if there are less than two active junctions the line will appear slightly thicker and red so as to indicate that the joint will be unstable. To make it return to a benign gray, make at least two junctions along the helical joint.

5.1.2 Points

Figure 14: Side View Plot Point Graphics

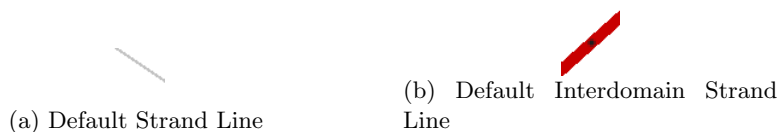


These symbols represent different artifacts, as follows...

Icon	Artifact
Triangle	NEMid (the midpoint between two nucleosides) whose direction matches that of the triangle.
Chevron	Nucleoside with an unset base whose direction matches that of the chevron.
Letter	Nucleoside with a set base that matches that of the letter, and whose direction is equal to that of the direction that the text reads in.
Circle	A nucleoside or NEMid that cannot be interacted with. When the Config Panel (4) is in Sequencing mode only nucleosides can be interacted with, for example, but NEMids are still displayed as small dots.

5.1.3 Lines

Figure 15: Side View Plot Line Graphics



The lines connecting dots together are not a literal representation of the location of the phosphate backbone of the DNA, but rather serve as a visual aid to see which strands are which. To this end, by default “interdomain” strands (strands that traverse multiple domains) are automatically styled to be thicker and more colorful to stand out; however, the styles of lines are completely customizable and need not be this way.

Additionally, for interdomain strands, the lines are curved so that it is clear which strand is which, because with the thicker lines, such as those which interdomain strands possess, strands appear to merge with themselves otherwise. At this time, this setting cannot be disabled.

5.2 Interaction

NATuG utilizes the pyqtgraph framework for its plots. Below lies a general description of the means of interaction for the plot, but, for a more direct and comprehensive breakdown, visit pyqtgraph’s website (pyqtgraph.readthedocs.io).

5.2.1 Panning

To pan from left to right within the side view plot, left click, and then drag your cursor away from the area that you want to pan to. You can think of this as “pulling” the graph towards the cursor. Alternatively, if you have a mouse, you also have the option of clicking the scroll wheel and dragging in the direction that you would like to move the plot in. If you wish to pan horizontally only you can click on the x axis labels and drag left/right, and likewise for vertical panning, but with the z-axis.

5.2.2 Zooming

To have the plot automatically zoom in out to showcase all of the items currently plotted, click the small “A” symbol in the bottom left of the plot. This is called the “Auto Range” button. For zooming in on specific areas of the plot, see below.

- If you have a mouse:
 - To zoom while maintaining the aspect ratio of the plot, use the scroll wheel.
 - To zoom without regard for the aspect ratio of the plot, right click, and then drag in the direction that you would like to stretch the plot in. The axes will automatically update.
- If you have a trackpad:
 - To zoom while maintaining the aspect ratio of the plot, pinch inwards to outwards to zoom in, and outwards to inwards to zoom out.
 - To zoom without regard for the aspect ratio of the plot, pinch and right click simultaneously.

Also it is worth noting that if you wish to zoom horizontally only you can click on the x axis labels and scroll, and likewise for vertical panning, but with the z-axis.

5.2.3 Configuration

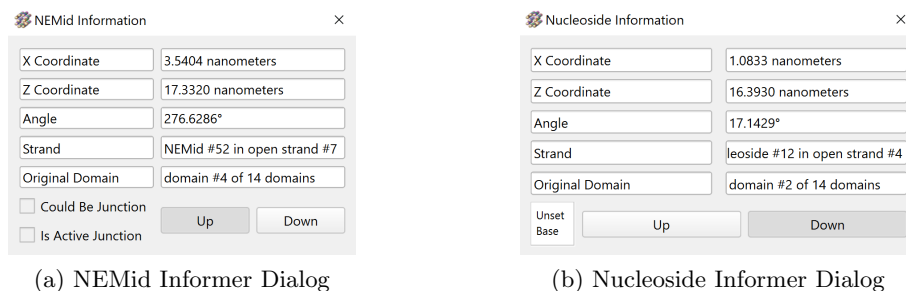
For additional options for the plot, right click. Upon right clicking, a small dialog showing additional plot options will show up. This allows for more advanced configuration of how the plot appears.

5.3 Modes

The currently chosen mode dictates what left clicks on non-dot points in the Side View Plot does. To change the mode, simply click on the mode that you would like to change to. Only one mode can be chosen at a time, and the currently chosen mode is indicated in a slightly darker gray than the others.

5.3.1 Informer Mode

Figure 16: Informer Dialogs



The “Informer” mode is for obtaining additional information about given points. It provides a dialog that shows various attributes of the point(s) that were clicked on.

The informer mode is for obtaining additional information about given points. When the informer mode is active, clicking on any point within the main plot will create a dialog that provides additional information about the point. If you click on a NEMid that could be made into a junction, two informers will appear—one for the NEMid that was clicked, and one for the NEMid that that NEMid could potentially be made into a junction with. There are two different types of dialogs: the NEMid Dialog ?? and the Nucleoside Dialog ??.

Nucleoside Informer Descriptions of the various properties that the Nucleoside Informer displays:

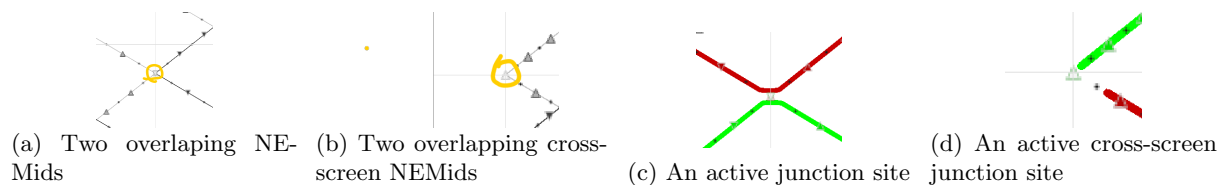
1. **X Coordinate:** The x coordinate of the nucleoside.
2. **Z Coordinate:** The z coordinate of the nucleoside.
3. **Angle:** The angle of the nucleoside. The line of tangency between the previous domain and this domain marks the zero degree mark. This is automatically moduloed by 360.
4. **Strand:** The index of the nucleoside within its parent strand, and the index of the parent strand within the list of all strands. Indexes begin at #1. This is the nucleoside index, not the item index, so even though it is true that strands progress nucleoside, NEMid, nucleoside, etc., this index only takes into account nucleosides.
5. **Original Domain:** The domain that this nucleoside belongs to.
6. **Base:** The base that this nucleoside is currently set to. If the nucleoside does not have a base set, this box will read "Unset Base."
7. **Direction:** The direction of the strand. The darker box is the direction that the strand is going in.

NEMid Informer Descriptions of the various properties that the NEMid Informer displays:

1. **X Coordinate:** Same as Nucleoside Informer #1
2. **Z Coordinate:** Same as Nucleoside Informer #2
3. **Angle:** Same as Nucleoside Informer #3
4. **Strand:** Same as Nucleoside Informer #4
5. **Original Domain:** Same as Nucleoside Informer #5
6. **Could Be Junction:** Whether this NEMid could be conjuncted with another NEMid. If this is set to true then clicking on this NEMid in Juncter Mode (5.3.2) will be allowed, otherwise an error will be displayed
7. **Is Active Junction:** Whether this NEMid is currently conjuncted with another NEMid.
8. **Direction:** Same as Nucleoside Informer #7

5.3.2 Juncter Mode

Figure 17: Junction Sites and Junctable Regions



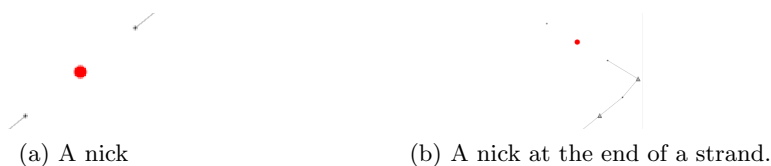
The "Juncter" mode allows for the creation of cross-strand exchanges. It lets you left click on overlapping NEMids to create junctions.

There are two ways to create a junction:

- Clicking on two overlapping NEMIds (indicated by white triangles that are on top of each other as shown in figure ??). This is the typical type of junction, and is highly versatile—it can split closed loops, form closed loops, allow a strand to go across many domains, and more. See figure ?? for an example of what a junction will look like after a left click.
- Clicking on a single NEMId that overlaps a NEMId on the other side of the screen (see figure ??). This type of junction is called a "cross-screen" junction, and can be seen in figure ?? . Since the Side View Plot is really an unrolled view of all of the helical domains, the very left for closed structures will sometimes have NEMId overlaps with the very right.

5.3.3 Nicker Mode

Figure 18: Nick examples



The "Nicker" mode is how you can create nicks within the strand. Nicks are essentially gaps that split a strand into two different strands.

To create a nick, click on any NEMId, and the NEMId will convert into a red circle. There will then be two strands: one consisting of all of the NEMIds and nucleosides that were above the NEMId, and one consisting of all the NEMIds and nucleosides below the NEMId. The NEMId itself will be removed.

Notes about nicking:

- When you create a nick the data of the original strand is saved, so, to remove a nick, click on the nick a second time and the original strand will return.
- Creating a nick at the end of a strand is unadvised because, as can be seen in figure ??, you will likely inadvertently create a strand that consists of a single nucleoside.
- The nitrogenous bases of all the nucleosides in the two new strands will be preserved.

5.3.4 Linker Mode

The "Linker" mode allows you to connect the end of one strand to the beginning of another strand, and vice versa.

Linkages A linkage is a single-stranded region that connects two formerly distinct strands. In this single-stranded region nucleosides can be added, and bases for those nucleosides can be set. To customize a linkage click on the linkage, and the linkage editor should pop up. The Linkage Editor (see figure ??) is a dialog that allows you to add nucleosides to the linkage, and to set the sequence of the items in the linkage. To increase or decrease the number of items within the linkage, simply increase the "Nucleosides in Linkage" spinbox, and click tab or enter to cause the update to take effect. Note that the nucleosides in the linkage will show up in the regular strand config sequence editor, but this window only shows the sequence of the items within the linkage. To remove a linkage, you'll need to open this same dialog and use the **remove linkage** button at the bottom.

To create a linkage:

1. Locate two NEMIds that can be made into a linkage. Consult the following list of prerequisites for determining if two NEMIds can be linked together—if the following conditions are not met a warning will be displayed.

Figure 19: The linkage editor

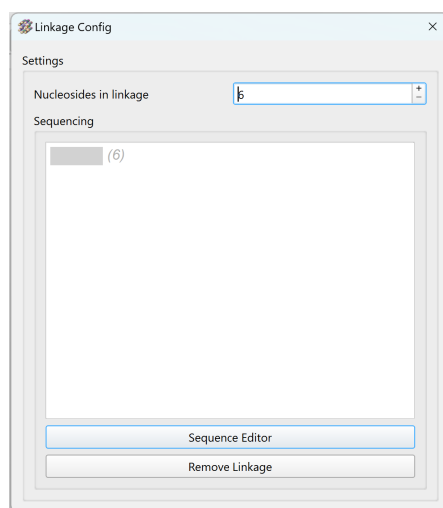
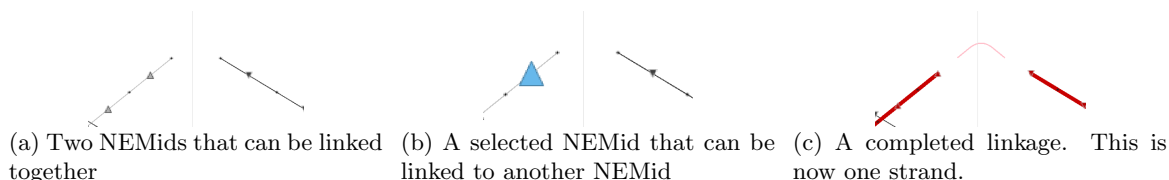


Figure 20: Linkage creation process



- (a) Both NEMIds must be endpoints of their respective strands. This means the second to last or second item in their strands (with the nucleoside that follows them being the very last or very first item). Recall that you can split up strands with nicks so as to 'create' endpoints.
 - (b) One NEMid is at the 5' end of its strand and the other NEMid is at the 3' end of its strand.
 - (c) Both strands consist of >1 NEMIds.
2. Left click on one of the NEMIds that you would like to link with the other NEMid. The order in which you click the NEMIds does not matter. After clicking on a NEMid in linker mode, the NEMid should turn blue and grow larger. This indicates that it is selected.
 3. Left click on the other NEMid that you would like to link it to. A linkage should be created upon releasing left click.

5.3.5 Highlighter Mode

The "Highlighter" mode lets you highlight points. It makes them larger and yellow, which can be useful for presentations (see figure ??). To highlight a NEMid or nucleoside, simply click on it. To unhighlight a NEMid or nucleoside, click on the highlighted NEMid or nucleoside a second time.

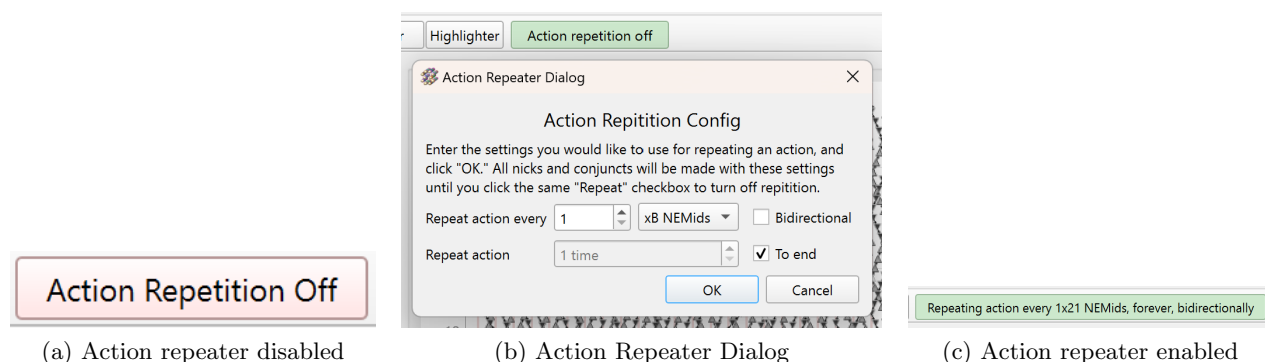
Notes about highlighting:

- Items will stay highlighted even when you leave highlighter mode. This means that there are actually multiple ways to unhighlight items, including, but not limited to the following.
 - Click the highlighted point while in highlighter mode



Figure 21: "A highlighted NEMid

Figure 22: Enabling the Action Repeater



- Create a nick to split the strand in two, or create junction
- Click the "Update Graphs" button to refresh the plots
- At this time NATuG's highlighting is uncustomizable, so all highlighted items will be large and yellow versions of their past selves.

5.4 Repetition

Often for larger structures, especially ones where helices contain many bases, you will find yourself tediously making junctions all the way up the helices every where and there. Thankfully, NATuG has a tool to streamline the process. With the Action Repeater, you can automatically repeat an action, such as creating junctions, along a **helix**, with the frequency of your choice. It is important to realize that the action repeater does not repeat actions across strands, but rather traverses up/down **helices**.

To enable the Action Repeater, click on the large, red button that says "Action Repetition Off," at the top of the screen. Clicking on it should spawn an Action Repetition Dialog, allowing you to configure and then enable the tool. The Action Repeater works for any mode, such as Nicking and Juncting, but if it fails to perform an action at any given point along the helix it is repeating over, it will suppress the error. Within the Action Repeater Dialog there are various settings you can configure, which should be mostly intuitive. You can choose to repeat the action every certain amount of NEMids, or every certain multiple of B (a configurable setting in the Nucleic Acid Profile) NEMids. You can also specify how far up the helix to go in those intervals, or can choose to continue to the end of the strand. Choosing the bidirectional checkbox will make it so that the action repeats up and down the strand (whereas normally it only repeats in the direction of the point you click, up until the end of the strand).

After enabling the Action Repeater, the large red button will turn green and display the current repetition settings. To disable the Action Repeater, simply click the large green button displaying the settings.

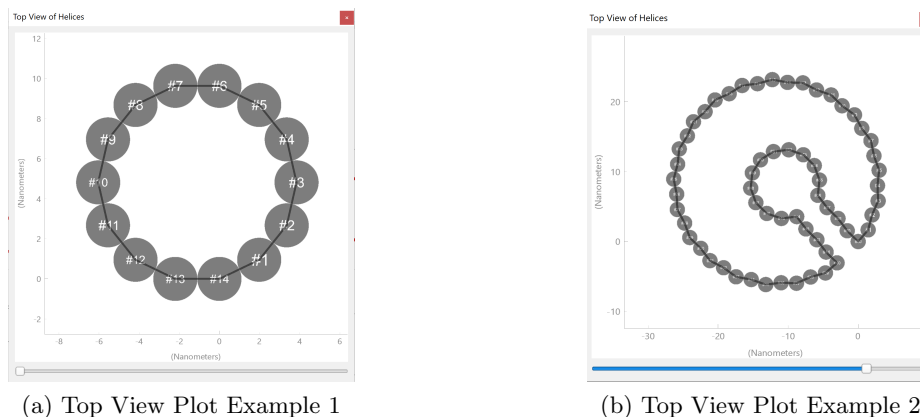
6 Top View Plot

The Top View Plot presents an overhead view of the nanostructure that is currently being created. It allows you to visualize the actual tube shape, and see the interior angles of all the domains. The Top View Plot is a panel, which means that it can be undocked, among other things. For information on panels, see ??.

6.1 Graphics

In the Top View Plot, each circle represents a helical domain, which is the region in which the double helices exist. The circles are numbered, and the numbers represent the domains' respective indexes. The units of the axes are in nanometers, and depend on the diameter of domains, which can be set in the Nucleic Acid Tab (4.2). An imaginary 0^{th} domain is drawn as well to show the last domain's interior angle, but this

Figure 23: Top View Plot Panel



domain does not actually exist. Additionally, to make it easier to see the interior angles of the domains, a thin line is drawn atop all of the domains' circles.

6.2 Interaction

For the most part, interacting with the Top View Plot is the same as interacting with the Side View Plot (documentation for interaction with the side view plot can be found at 5.2). However, there are a few important differences:

- The aspect ratio of the Top View Plot is locked at 1:1. This means that you cannot stretch the plot.
- The Top View Plot is rotatable. To rotate the top view plot, drag the handlebar beneath the plot to the right. The left edge of the bar represents a 0° rotation, and the right edge of the bar represents a 360° rotation. The plot is rotated after coordinates for all of the domains are computed, and the pivot point of rotation is the middle point of all the domains.
- Clicking on a number (not a circle, but rather on a number) will navigate the Side View Plot to the specific domain that was clicked on. Clicking on the same number a second time will restore the full view of the plot.

7 Saving

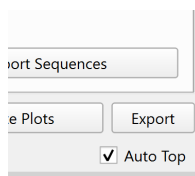
A vital feature of NATuG is the ability to save and load the current program state at any time. NATuG makes this extremely simple: either use **Control+S**, or go to file, save. Your save will take the form of a **.natug** file, which contains the entire program state, with the exception of current snapshots, which are their own save files and are not packed into the save. Technically speaking, NATuG save files are **zip** formatted packages, so, while not documented here, all program contents can be found in neat **.csvs** within the archive.

8 Exporting

One important feature of NATuG is the ability to export the side view plot as a high resolution vector or image file. These images are of a high enough fidelity to be used for scholarly literature and also allow for a greater level of image customization.

To begin the export process, click the “Export plots” button on the bottom right of the Config Panel. After you have done so, a Plot Exporter dialog will pop up that is fixed at 85% of your screen’s width and 80% of your screen’s height.

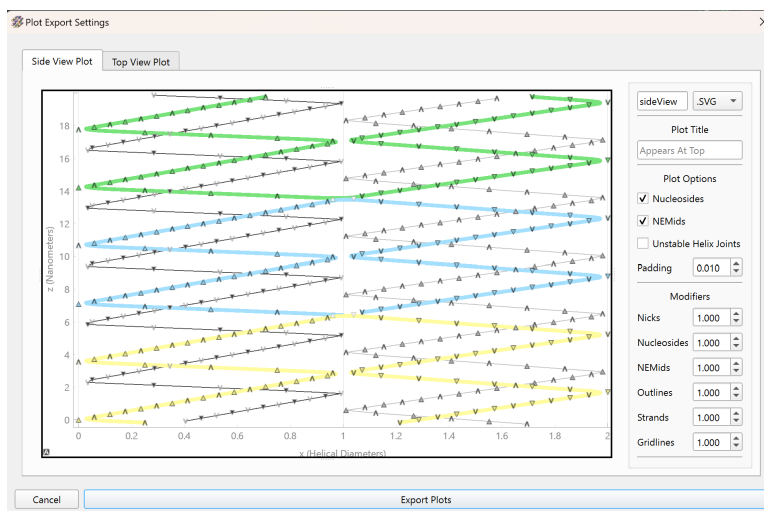
Figure 24: Export plots button



The Plot Exporter dialog has two primary sections. On the left, there is a tabbed area with two tabs for each of NATuG's two primary plots: a Side View tab and a Top View tab. On the right side there is a configuration panel that allows you to customize the currently displayed plot. Changes made within the settings area are saved as you switch back and forth between tabs.

8.1 Export Settings

Figure 25: Plot Export Dialog



On the right side of the Plot Exporter dialog you can configure how the plots generate. Settings like the title displayed above the plot, the size of the various artifacts, and the padding around the plot can be configured easily. Below is a description of the various properties that can be configured for the Side View Plot and the Top View Plot.

In the **primary settings** export options area the export filetype can be set within a dropdown to be one of the following:

- **SVG:** This is the recommended filetype, and is saves the plot as a vector graphic. That means that the image is of infinite resolution, and can be edited in vector editing programs like Adobe Illustrator later.
- **PNG:** A PNG image file.
- **JPG:** A JPG image file.

The Side View Plot and Top View Plot both have similarly categorized settings, but the options themselves differ. The following subsections explain how the different settings affect the way that the plots are generated.

8.1.1 Side View Plot Settings

For the Side View Plot, there's different types of controls that you have over the plot:

- **Primary Settings:**
 - **File Name:** The name of the exported file.
 - **Plot Title:** The title to place at the top of the plot contents. If blank no plot title is plotted.
- **Plot Options:** General properties apropos how to plot the artifacts, and which to plot. This includes:
 - **Nucleosides & NEMids:** Whether to plot nucleosides or NEMids. Plots them if checked, does not if they are unchecked.
 - **Unstable Helix Joints:** Whether to make the vertical gridlines between domains slightly wider and red if there are less than two active junctions along the helix joint. If this is checked the gridline is altered, otherwise it is gray like the others.
 - **Padding:** When autoranged (the small “A” autorange button on the bottom left of the plot is clicked and the plot contents then automatically are rescaled to fit the plot), this is the amount of padding to place around the plot. When a manual range on the plot is chosen this property does not have any effect.
- **Modifiers:** Multipliers for the size of various plot artifacts. Artifacts' default sizes (or widths, when applicable) are multiplied by these values.

8.1.2 Top View Plot Settings

- **Primary Settings:**
 - **File Name:** The name of the exported file. If this is the same as that of the Side View Plot then the Top View Plot export will overwrite the Side View Plot's and only one file will end up being created.
 - **Plot Title:** The title to place at the top of the plot contents, like that of the Side View Plot. Likewise, if blank no plot title is plotted.
- **Plot Options:** General properties apropos how to plot the artifacts, and which to plot, similar in essence of those of the Side View Plot's. These include:
 - **Rotation:** How much to rotate the plot contents, in degrees.
 - **Stroke:** The width of the stroke of the circles.
 - **Padding:** The same as the Side View Plot's padding.

9 Strands

Whereas a “helix” is specifically a strand that stays within one domain, a “strand” can traverse many domains. When a strand crosses through many domains, it weaves the nanotube together. But, equally important to choosing where the strand goes is determining the base sequence. And, being able to choose styles to represent strands in publications is also critical. This section covers how to customize the properties of a strand.

9.1 Selecting a Strand

To select a strand, left click on the line between two NEMids/nucleosides within the Side View Plot. The Strand Config Dialog should pop up, and the NEMids/nucleosides of the strand should grow larger (this is because the width/color of the strand is customizable, so it would be annoying if the strand's color/width changed when highlighting). Only one strand can be selected at a time. Multiple strands can be selected at a time, and the name of the window indicates the strand that the Strand Config Dialog is for.

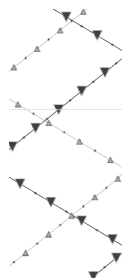


Figure 26: A selected strand

9.2 Configuration

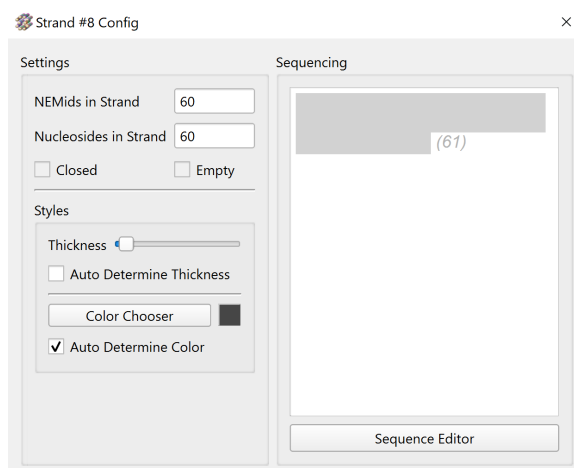


Figure 27: Strand Config Dialog

The Strand Config Dialog is the place where you can configure strand styles, obtain additional information about a given strand, and set a strand's sequence. Below there is a detailed description of the three main parts of the strand config dialog:

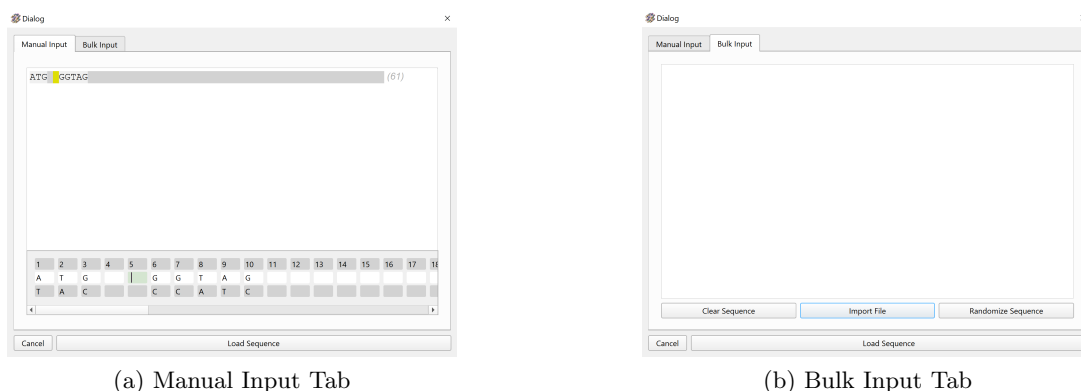
Settings The settings area of the Strand Config Dialog encompasses the entire left side of the dialog. Within this area, information about the strand can be found, and various styles can be customized.

Information The information area of the Strand Config Dialog provides the following information about the strand:

- **NEMids in Strand:** The number of NEMids that are in the strand.
- **Nucleosides in Strand:** The number of nucleosides that are in the strand.
- **Closed:** Whether the strand is a closed loop strand.
- **Empty:** Whether the strand has no items in it. In general, this will be false.

In addition, the strand's index in regards to all the other strands (its "id") can be found as the title of the window (see the top of figure ??)

Figure 28: Sequence Selector Dialog



Styling The style area of the Strand Config Dialog lets you customize styles of the strand. The styles update in real time, with the exception of the “Automatic” style checkboxes. The “Auto” checkboxes next to the different strand styles tell NATuG whether or not to overright the user-set strand style when refreshing/recoloring the plot. If you want to reset a style of a strand back to the default, select the “Auto” checkbox for the style, and then in the file menu choose “View,” and “Restyle Strands.”

Below lies a list of the various styles that can be customized, and the default settings of the given styles.

- **Thickness:** The width of the strand. This is a handlebar that represents the width of the strand in pixels. The very left of the bar represents 1 pixel of thickness, and the very right of the bar represents 50 pixels of thickness. By default, strands that stay in their own domain are 2 pixels thick, and strands that traverse multiple domains are 9.5 pixels thick.
- **Color:** The color of the strand. The “Color Chooser” button creates a Color Chooser Dialog that allows you to select the color of your choosing. Within the dialog, you can enter parameters for your color in RGB/CMYK, choose a screen color, or choose from a preset. By default, strands that stay in their own domain are either light or dark gray, and strands that traverse multiple domains are automatically assigned colors as to keep them distinct from other strands that traverse multiple domains.

9.3 Sequence Selector

The purpose of the Sequence Selector is to provide a user-friendly way to obtain a valid sequence for a given strand/portion of a strand. Generally, this will pop up when you click a button along the lines of “Choose a Sequence.” To use the Sequence Selector, you can either choose to use the Bulk Sequence Input tab, or, if the length of the sequence you are being prompted to choose is under 1,000 bases, the Manual Input Tab. Once you have made your selection of sequence, click the “Load Sequence” button, or click the “Cancel” button to exit the Sequence Selector and cancel the sequence choosing operation. Below, more information about the two areas of the Sequence Selector can be found.

Note: When you set a sequence, whether through this selector dialog or elsewhere, complementary bases will automatically be set for you.

9.3.1 Manual Input Tab

The Manual Input Tab of the Sequence Selector is like a text editor for a DNA sequence, and makes entering a sequence manually super-simple and convenient. Because of how the Manual Input Tab is implemented, it becomes particularly slow after displaying more than 1,000 bases, so, when more than 1,000 bases need to be set you must use the Bulk Input Tab. The Manual Input tab consists of two main areas, the bottom Sequence Entry Area and the top Sequence Display Area.

Sequence Entry Area The Sequence Entry Area is a horizontally scroll-able area that showcases all of the bases currently chosen as editable text boxes. Each white rounded box, also called an Entry Box, represents a single nitrogenous base. The box above each Entry Box is the index of that base. The box below each Entry Box is the complementary base. NATuG uses Watson-Crick base pairing, so the complement of A is T, the complement of T is A, the complement of G is C, the complement of C is G, and the complement of None is None.

When you are using the Sequence Entry Area, you can type into one box at a time either the letter “A,” “G,” “C,” or “T,” or you can click the delete key. What you type will automatically be capitalized, and after you enter the letter of a base you will automatically be shifted to the next Entry Box to the right, or, if you enter a base into the last box, you will be shifted to the first Entry Box. As you enter bases into the various Entry Boxes, the Sequence Entry Area will automatically scroll to keep the currently selected box in view.

Sequence Display Area The Sequence Display Area is a larger area that displays the entire current sequence. This area is read only, and highlights the currently selected Entry Box of the Sequence Entry Area. You can select and copy the text of this box, but you cannot edit this box.

9.3.2 Bulk Input Tab

The Bulk Input Tab of the Sequence Selector is a place to load in or generate a large number of bases easily. Within the Bulk Input Tab there is only a **display area**, wherein there are currently two ways to populate it with bases:

1. Import a `.txt` file full of bases. The file must only contain “A,” “G,” “C,” or “T”s, along with linebreaks and spaces if desired (but whitespace has no effect on the load-in process). The example `m13mp18` virus strand sequence ships with NATuG as an example, and, as you will come across with your own viral sequences too, it is probably going to include more bases than the length of your longest strand. When you attempt to load it a warning will be displayed, which will allow you to apply the bases from the start of the file up until the strand you are setting a sequence for has been filled.
2. Generate a random string of “A,” “G,” “C,” and “T”s that is of the length of the strand.

Once you’ve imported a longer sequence with the bulk input tab, if it is under 1,000 bases you can always click “load sequence,” and go back to the Sequence Selector’s Manual Input Tab to fine-tune.