SUPPORTING INFORMATION

Characterization of Biomolecular Helices and Their Complementarity Using Geometric Analysis

Kevin Hauser[1], Yiqing He[2], Miguel Garcia-Diaz[3], Carlos Simmerling[1,4,*] and Evangelos Coutsias[4,5,*]

[1]Department of Chemistry, Stony Brook University, Stony Brook, New York, 11794
[2]Great Neck South High School, Great Neck, New York, 11023
[3]Department of Pharmacological Sciences, Stony Brook University, Stony Brook, New York, 11794
[4]Laufer Center for Physical and Quantitative Biology, Stony Brook University, Stony Brook, New York, 11794
[5]Department of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, New York, 11794
*Corresponding address: E.: carlos.simmerling@stonybrook.edu, E.: evangelos.coutsias@stonybrook.edu

Table S1. Cartesian coordinates of Nievergelt's helix.

x	y	z
12	102	198
48	138	180
65	163	169
77	187	157
85	209	149
94	266	128
93	288	120
89	316	112
82	347	107
62	397	103

Table S2. Residues whose Cα atoms were used to define the BurrH superhelix.

Module	Residue
1	Q42
2	P75
3	S108
4	S141
5	P174
6	P207
7	P240
8	P273
9	P306
10	P339
11	P372
12	P405
13	P438
14	L471
15	L504
16	L537
17	R570
18	A603
19	A636
20	P669
21	P702
22	P734
23	P765

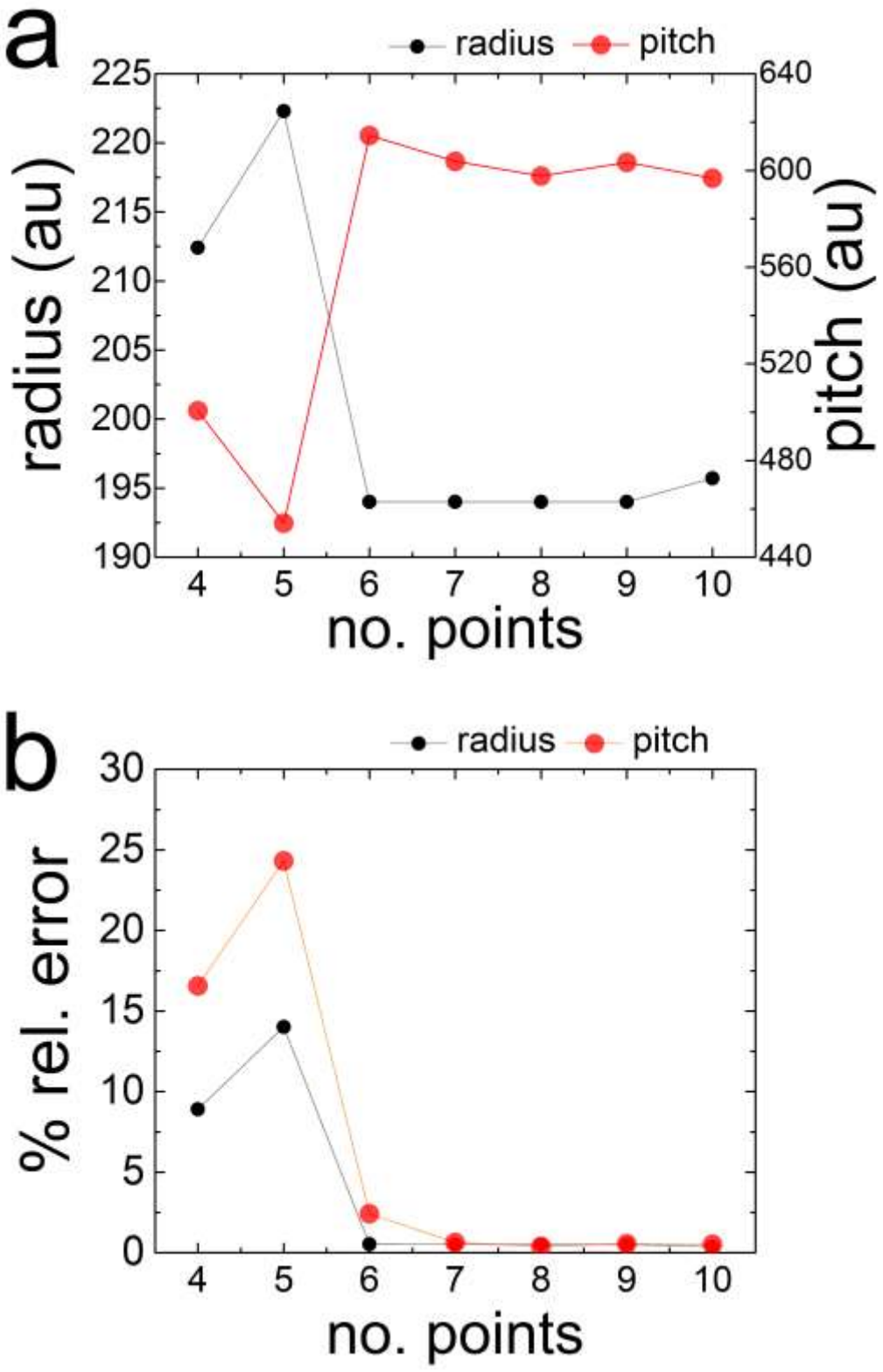

Figure S1. Helix parameters of Nievergelt's helix. (**a**) Helix pitch and radius (right and left Y axis respectively), with increasing number of points (X axis). (**b**) Percent relative error (Y axis), using Nievergelt's data as the reference, with increasing number of points (X axis).

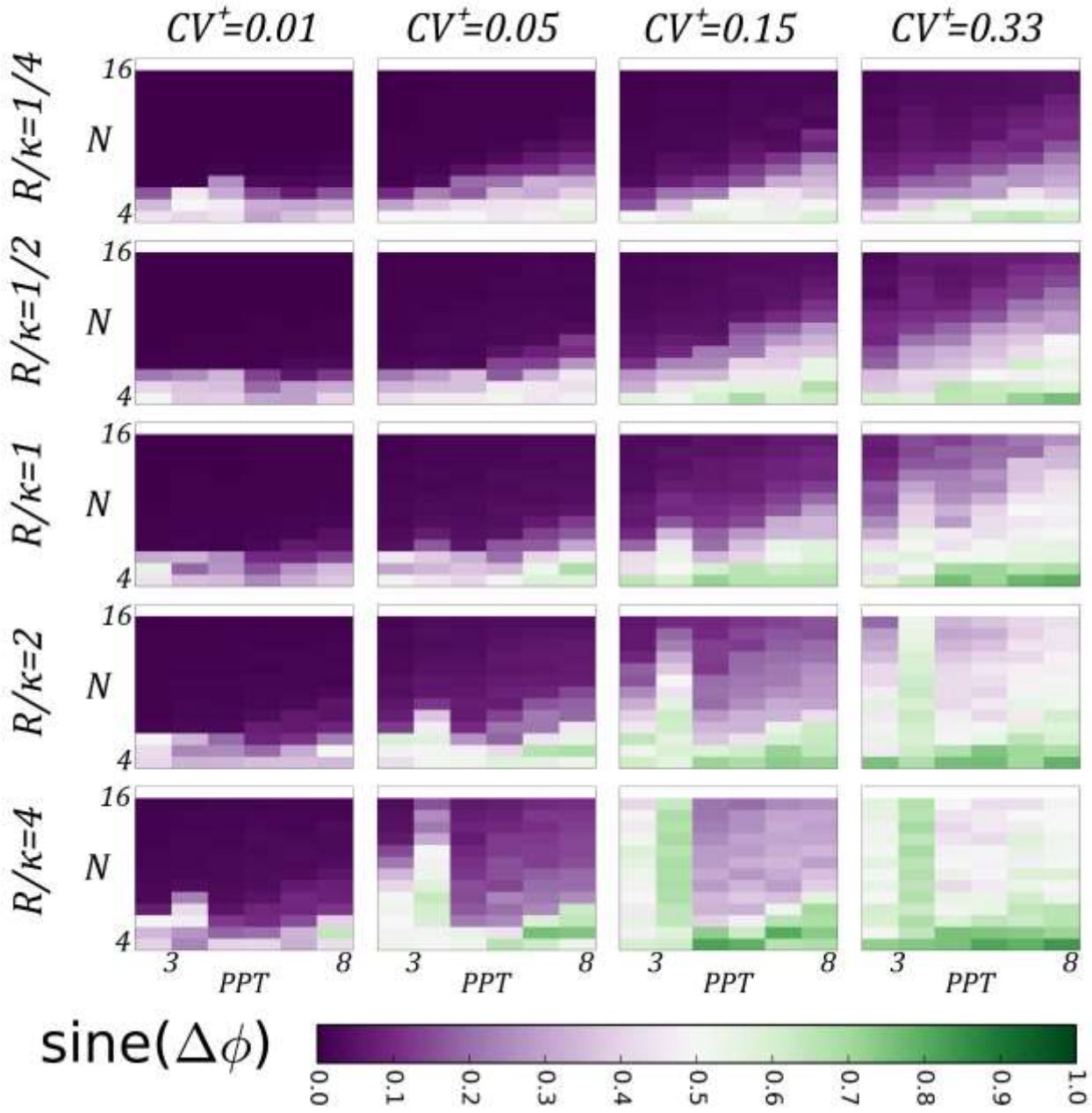

Figure S2. Sine of the average difference in the helical axes of the test helices and the ideal helices. Helix matrices are organized per **Figure 3**. We plot the sine of the difference in the average spherical coordinate φ-angle of the optimal helical axes for the 256 noisy helices in each cell and the spherical coordinate φ-angle of the optimal helical axis from the corresponding ideal helix If the difference is 0°, then sineΔφ is 0; if the difference is 45°, then sineΔφ is 0.5; if the difference is 90°, then sineΔφ is 1.

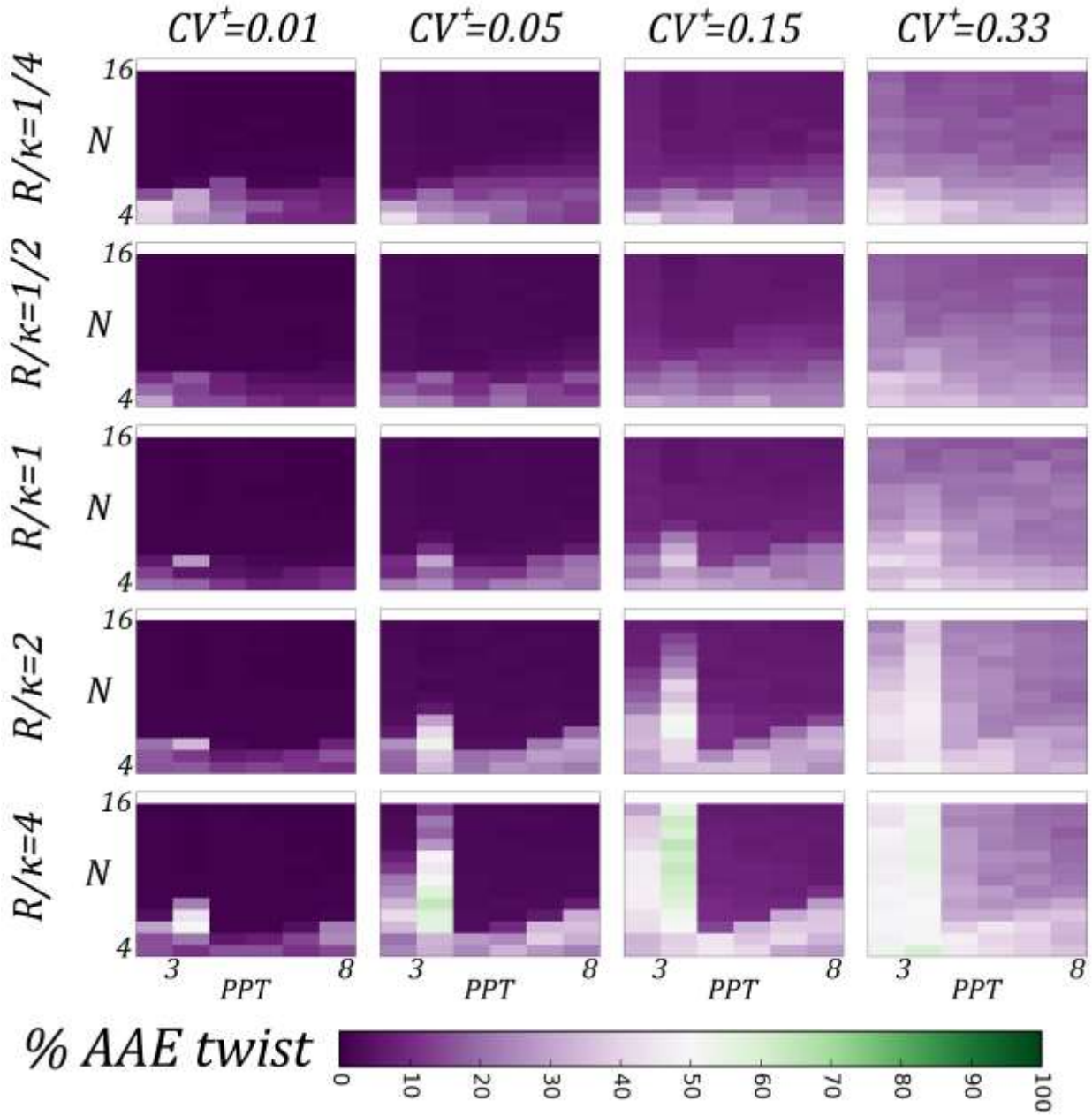

Figure S3. Percent AAE of derived helix twist for the 399,360 test helices. Helix matrices are organized per **Figure 3**. The plotted range of the percent average absolute error is 0 to 100.

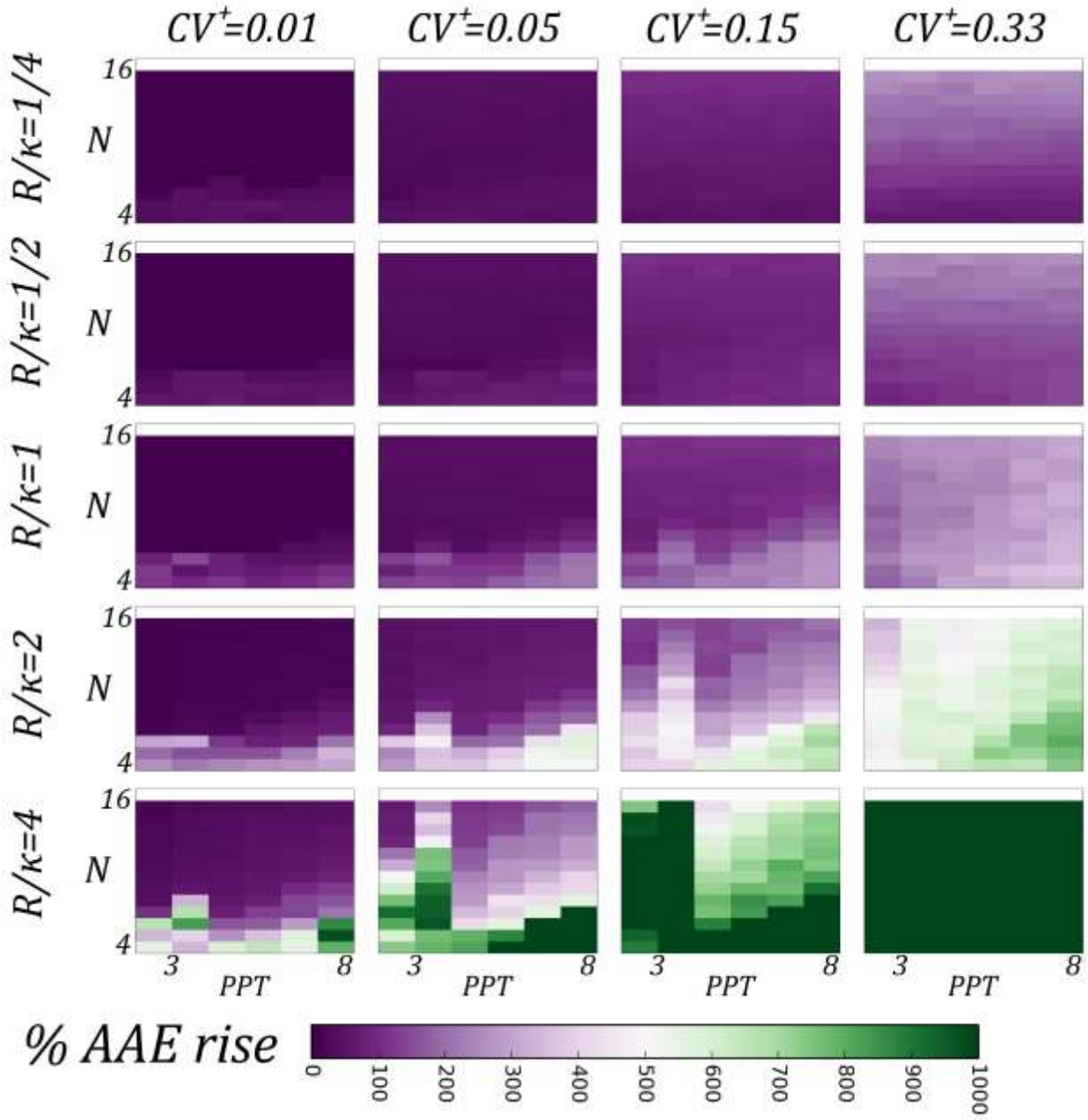

Figure S4. Percent AAE of derived helix rise for the 399,360 test helices. Helix matrices are organized per **Figure 3**. The plotted range of the percent average absolute error is 0 to 1000.

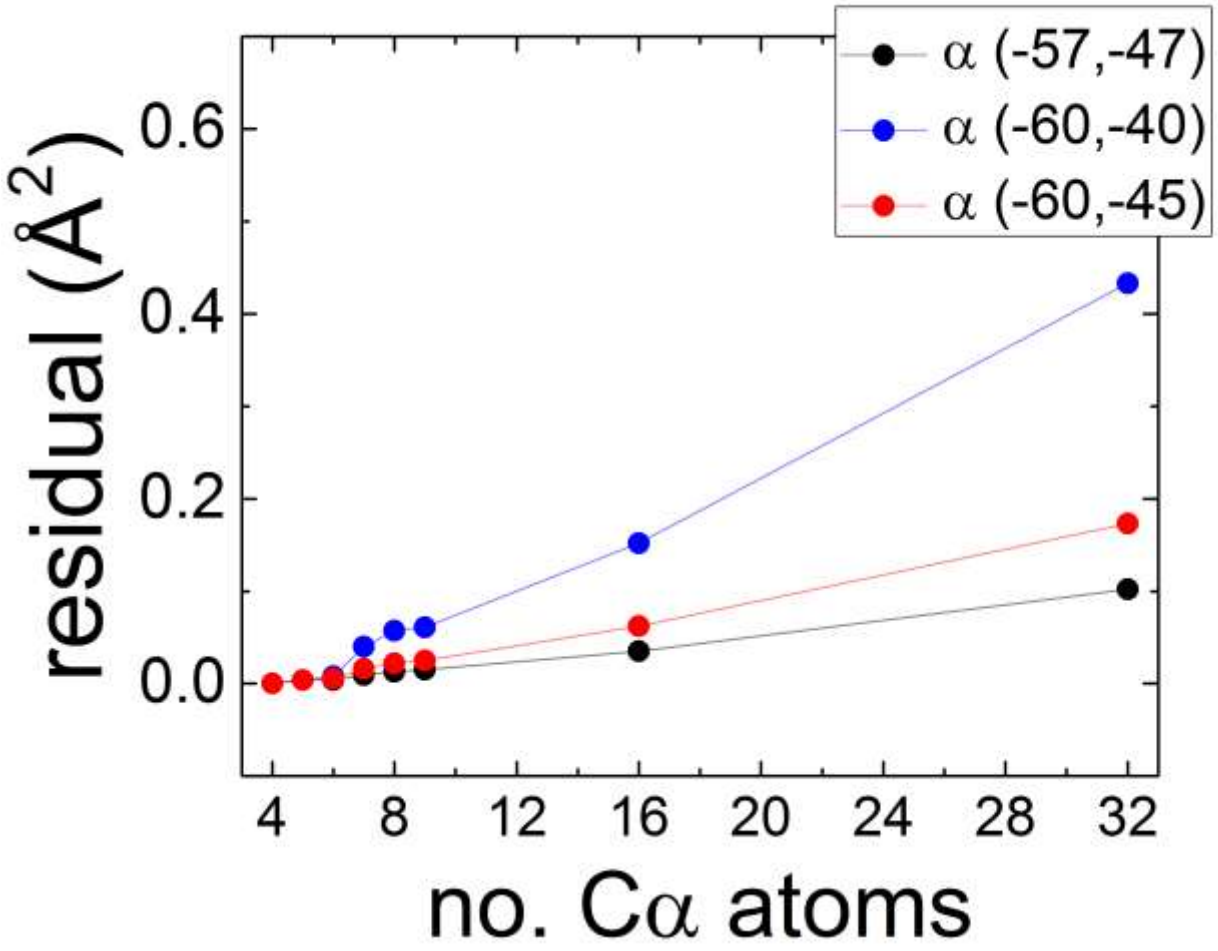

Figure S5. Fitting residual of three slightly different α-helical peptide secondary structure elements. Fitting residual was calculated using equation (12) and is plotted on the Y axis. The number of Cα atoms (one per amino acid) used in the fitting is plotted on the X axis. Residual increases with the number of Cα atoms used in the fitting because each atom contributes to the total deviation; thus, residual rises nearly linearly with number of atoms.

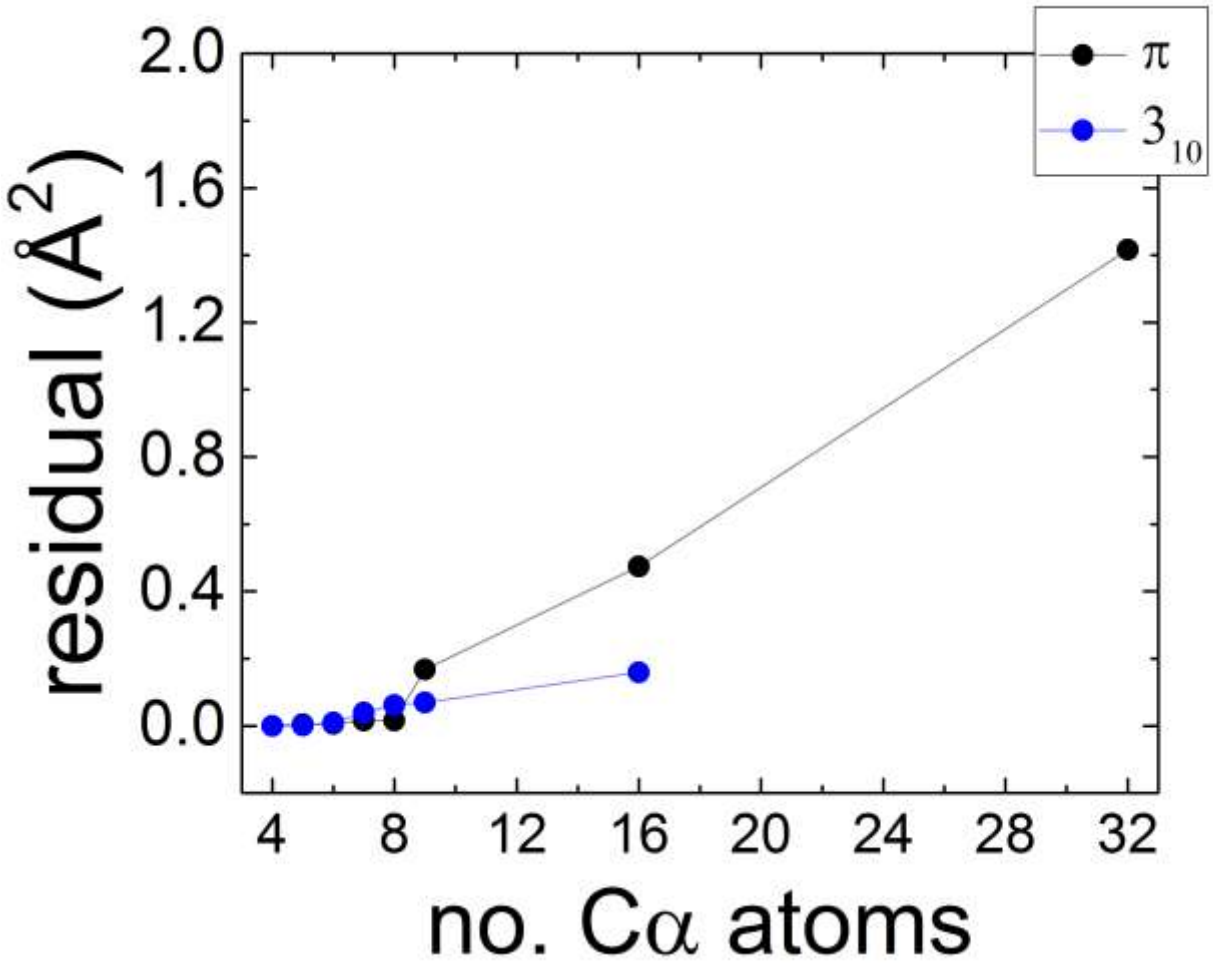

Figure S6. Fitting residual of π- and 3_{10}-helical peptide secondary structure elements. Fitting residual was calculated using equation (12) and is plotted on the Y axis. The number of Cα atoms (one per amino acid) used in the fitting is plotted on the X axis. Residual rises with the number of Cα atoms used in the fitting because each atom contributes to the total deviation; residual increases nearly linearly with number of atoms.

Table S3. RNA and DNA rise and twist are accurately calculated compared with Curves+[1] and 3DNA[2].

Nucleic acid	**Helix property**	**Curves+**[1]	**3DNA**[2]	**Helios (our method)**
A-DNA	rise (Å)	2.55 ± 0.00	2.54_8	2.56 ± 0.02
	twist (°)	32.7 ± 0.0	32.7	32.6 ± 0.1
B-DNA	rise (Å)	3.37 ± 0.00	3.37_5	3.37 ± 0.02
	twist (°)	36.0 ± 0.0	36.0	36.1 ± 0.1
A-RNA	rise (Å)	2.81 ± 0.00	2.81_2	2.80 ± 0.02
	twist (°)	32.7 ± 0.0	32.7	32.5 ± 0.1

All three nucleic acid structures were built using 3DNA[2]. For A-DNA and A-RNA, 11-bp duplexes were used. For B-DNA, a 10-bp duplex was used. One extra digit is provided for rise parameters of 3DNA because such was the precision used to generate the coordinates.

Software

The software associated with this method is available for download as part of the supplementary material attached with this article. The filename is `kHelios.tar.cf`. Once the file has been downloaded and moved to a folder, untar the file (using a terminal):

```
tar -xvf kHelios.tar.cf
```

A Fortran compiler is needed to install (compile) the software. In the example that follows, it is assumed the compiler is `gfortran` (which is free). If an alternate fortran compiler is desired, the installation script (see below) must be modified by replacing "gfortran" appropriately).

In the terminal, move into the `kHelios` directory that was created in the above step:

```
cd kHelios
```

Finally, install the software using the installation script `install_helios.sh`:

```
./install_helios.sh
```

Follow the on-screen instructions. Test cases (sample use of the software) are provided in the `kHelios/test_cases` directory.

References cited

1. Lavery, R.; Moakher, M.; Maddocks, J. H.; Petkeviciute, D.; Zakrzewska, K., Conformational Analysis of Nucleic Acids Revisited: Curves+. *Nucleic Acids Res.* **2009,** *37* (17), 5917-29.
2. Zheng, G.; Lu, X.-J.; Olson, W. K., Web 3DNA—A Web Server for the Analysis, Reconstruction, and Visualization of Three-Dimensional Nucleic-Acid Structures. *Nucleic Acids Res.* **2009**, gkp358.
3. Lawson, C. L. H., Richard J, *Solving Least Squares Problems*. SIAM: Philadelphia, 1995.