Supplementary Material

Lone pair…π contacts and structure signatures of r(UNCG) tetraloops, Z-turns and Z-steps: A WebFR3D survey

Craig L. Zirbel[1] and Pascal Auffinger[2*]

[1.] Department of Mathematics and Statistics, Bowling Green State University, Bowling Green, OH 43403, USA;
[2.] Architecture et Réactivité de l'ARN, UPR 9002, Université de Strasbourg, Institut de Biologie Moléculaire et Cellulaire du CNRS, 67084 Strasbourg, France;

Tel: +33 388417049; Fax: +33 388602218;
Email: p.auffinger@ibmc-cnrs.unistra.fr;

Contents:

Figure S1: A WebFR3D symbolic query for finding all lp…π contacts.

Figure S2: A WebFR3D symbolic query for finding NpN Z-steps.

Figure S3: A strict symbolic query for refining Z-turn searches with a 2.0 Å resolution limit.

Figure S4: WebFR3D finds GNNA Z-turns.

Figure S5: WebFR3D finds Z_{anti}-turns that fold similarly to Z-turns.

Figure S6: WebFR3D hits as visualized over the four nucleobases with a sO search at 3.0 Å resolution.

References

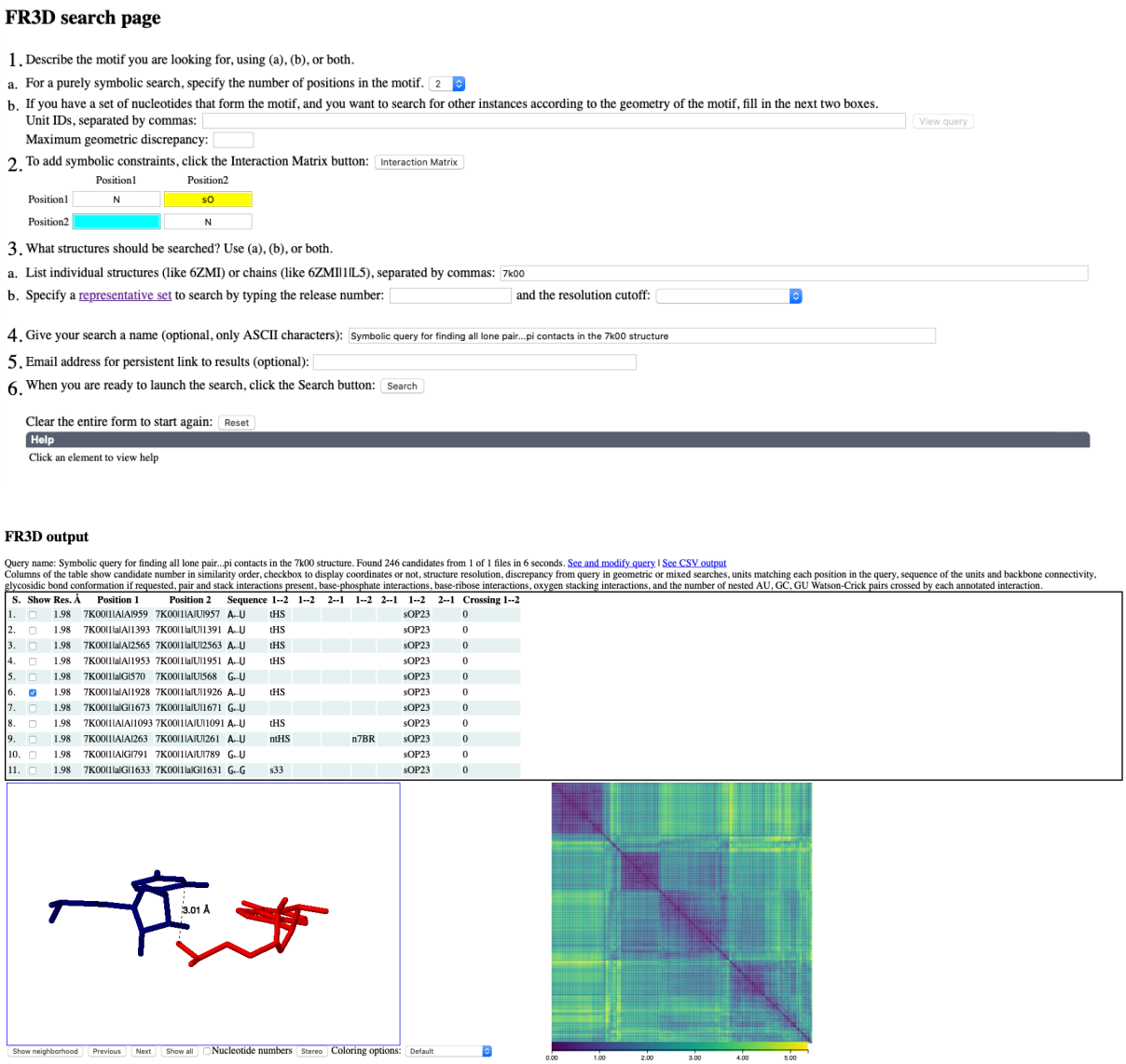

Figure S1: A WebFR3D symbolic query for finding all lp…π contacts. (Top) The search panel using the text string "sO" retrieves all instances of a stacking of a backbone oxygen atom (OP1, OP2, O2', O3', O4', and O5') from nt1 with a nucleobase of nt2. The string "sO" expands to "sO3 sO5" which expands to include all backbone oxygens. (Bottom) Partial view of the WebFR3D result panel showing a sOP23 contact (OP2 atom stacking with the 3'-face of a uridine) present in the 7K00 *Escherichia coli* 70S cryo-EM structure [1]. The shortest lp…π contact visualized by a dashed line involves the (U)C2 atom. Note that 246 potential lp…π contacts are identified in 7K00. Use the following link to retrieve the results of this search:
http://rna.bgsu.edu/webfr3d/Results/6287e7c307b40/6287e7c307b40.html

(last accessed the 6th of July 2022)

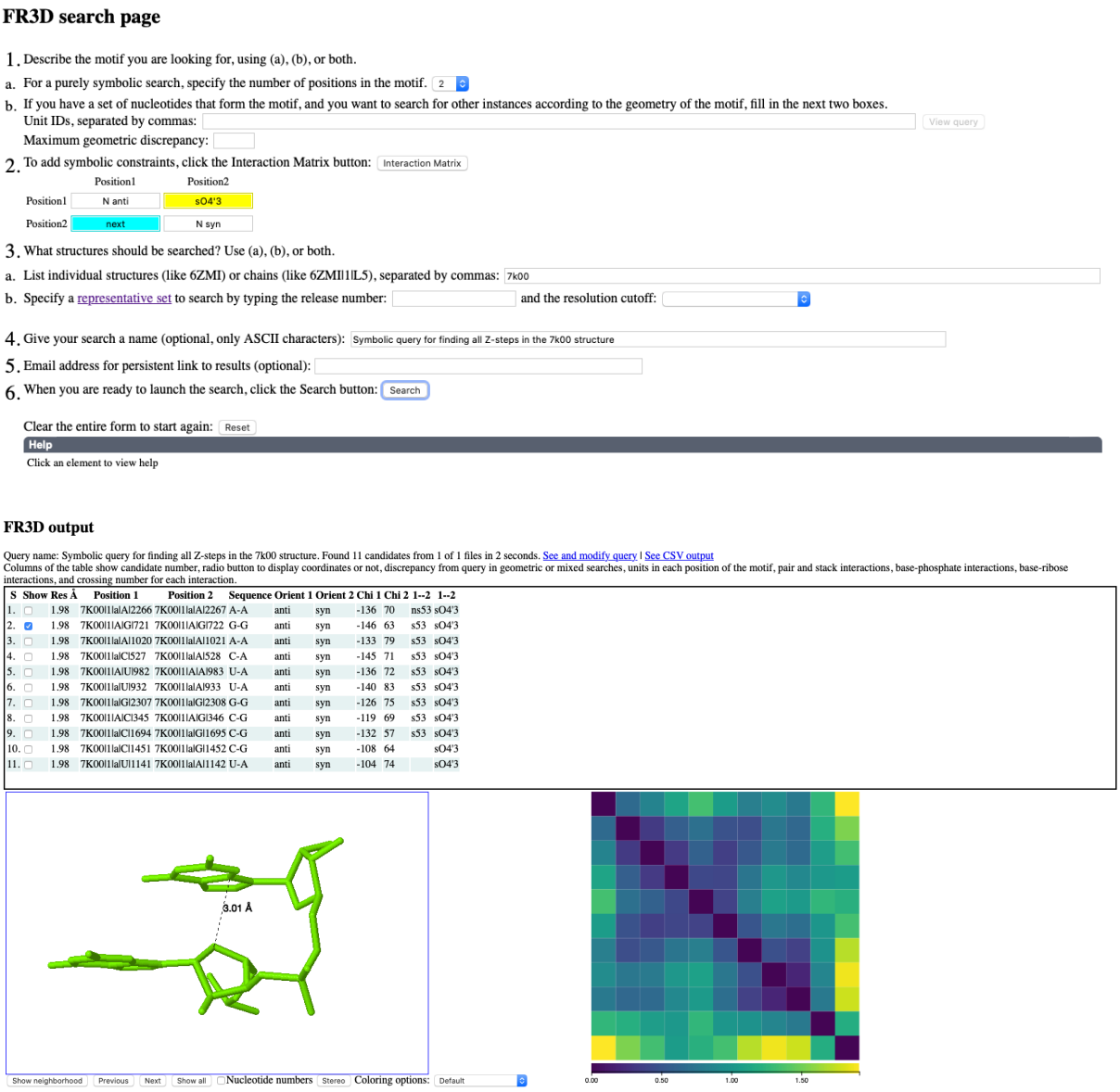

Figure S2: A WebFR3D symbolic query for finding NpN Z-steps. (Top) The search panel using the text string (sO4'3) indicates that all instances of a stacking of a O4' atom with the 3'-face of a nucleobase will be retrieved. To gather Z-steps, it is also necessary to specify that the nucleotides are consecutive ("next" condition), that the O4' ribose of the 5'-nt stacks with the 3'-face of the 3'-nt that is in *syn*. In a few instances, the 3'-nt can also be *anti*. When this constraint is given the search retrieves Z_{anti}-steps [2]. (Bottom) Partial view of the WebFR3D result panel showing a GpG Z-step in the 7K00 structure (note the head-to-head orientation of the riboses; see also **Figure 3b**). The lp…π shortest contact distance that is visualized by a dashed line involves the (G)C4 atom. Note that eleven potential NpN Z-steps are identified in the 7K00 *Escherichia coli* 70S structure. Use the following link to retrieve the results of this search:
http://rna.bgsu.edu/webfr3d/Results/627e8a7b5856f/627e8a7b5856f.html

(last accessed the 6th of July 2022)

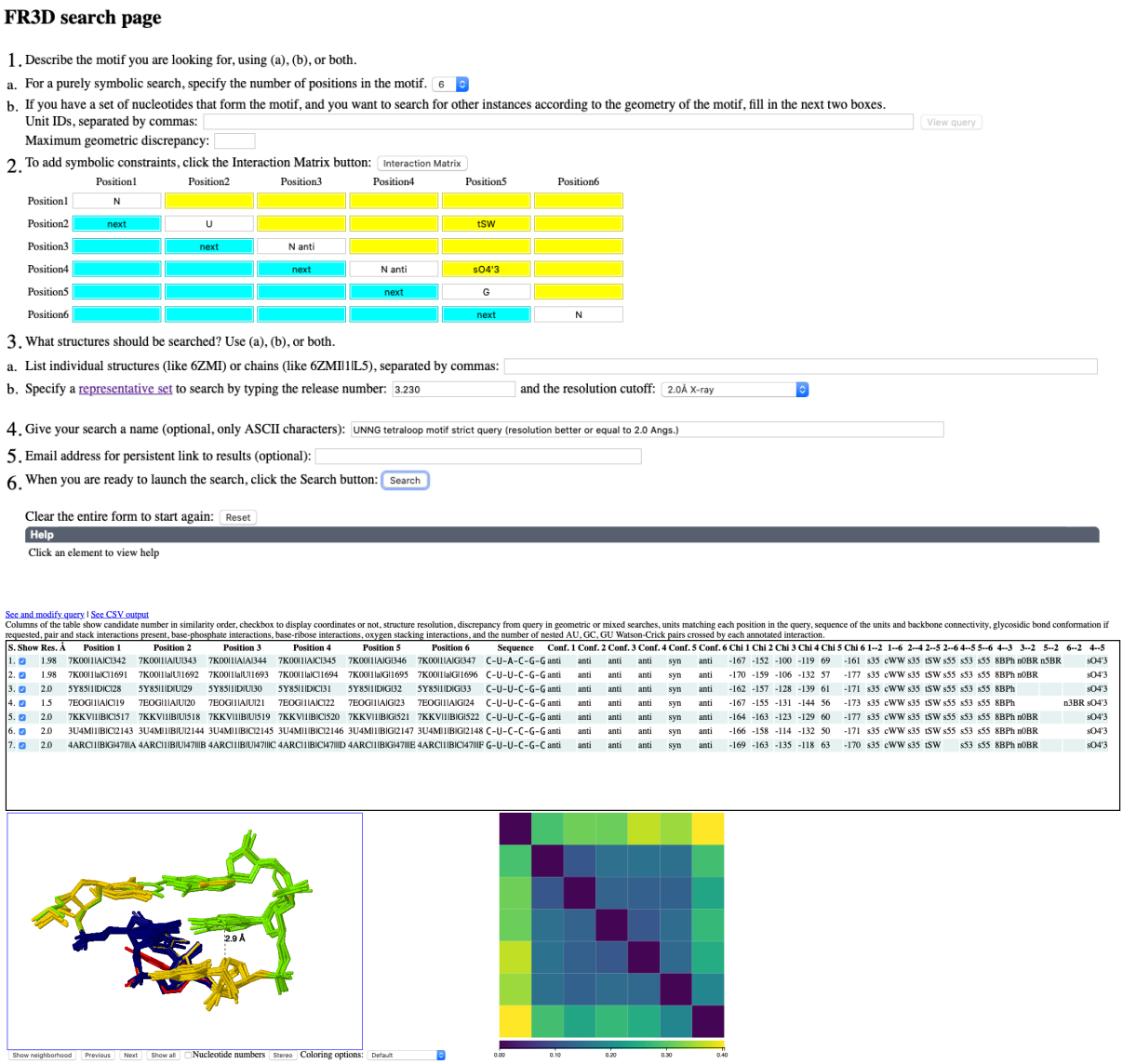

Figure S3 A strict symbolic query for refining Z-turn searches with a 2.0 Å resolution limit. (Top) WebFR3D search page with more stringent search criteria than those used in **Figure 5.** (Bottom) Partial view of the WebFR3D result panel showing an overlay of seven structures. The 2.9 Å lp…π contact shown in **Figure 5** for 3U4M is visualized by a dashed line. The rationales for applying more stringent search criteria are described next. By examining the result panel of **Figure 5**, we noted that four of the 14 turns do not display strict Z-step lp…π contacts. Therefore, we added a nt4…nt5 sO4'3 constraint. This narrows the search down to ten hits. Two additional Z-turns display a s33 stacking between nt4 and nt5 instead of the more common r(UNNG) s53 stacking pattern. This reveals that the bottom nucleotide is in *syn* and not *anti*. By discarding these two structures through imposing an *anti*-conformation to the 3rd tetraloop residue, the structural ensemble comprises now eight elements. A further examination of the **Figure 5** data reveals that the 2nd residue of 6DCB is in "Intermediate *syn*" and as noted in **Section 3.4** of the main text, the local electron density associated with this tetraloop is of poor quality motivating the 6DCB exclusion from the set of structures. Use the following link to retrieve the results of this search:
http://rna.bgsu.edu/webfr3d/Results/6287e67695a70/6287e67695a70.html

(last accessed the 6th of July 2022)

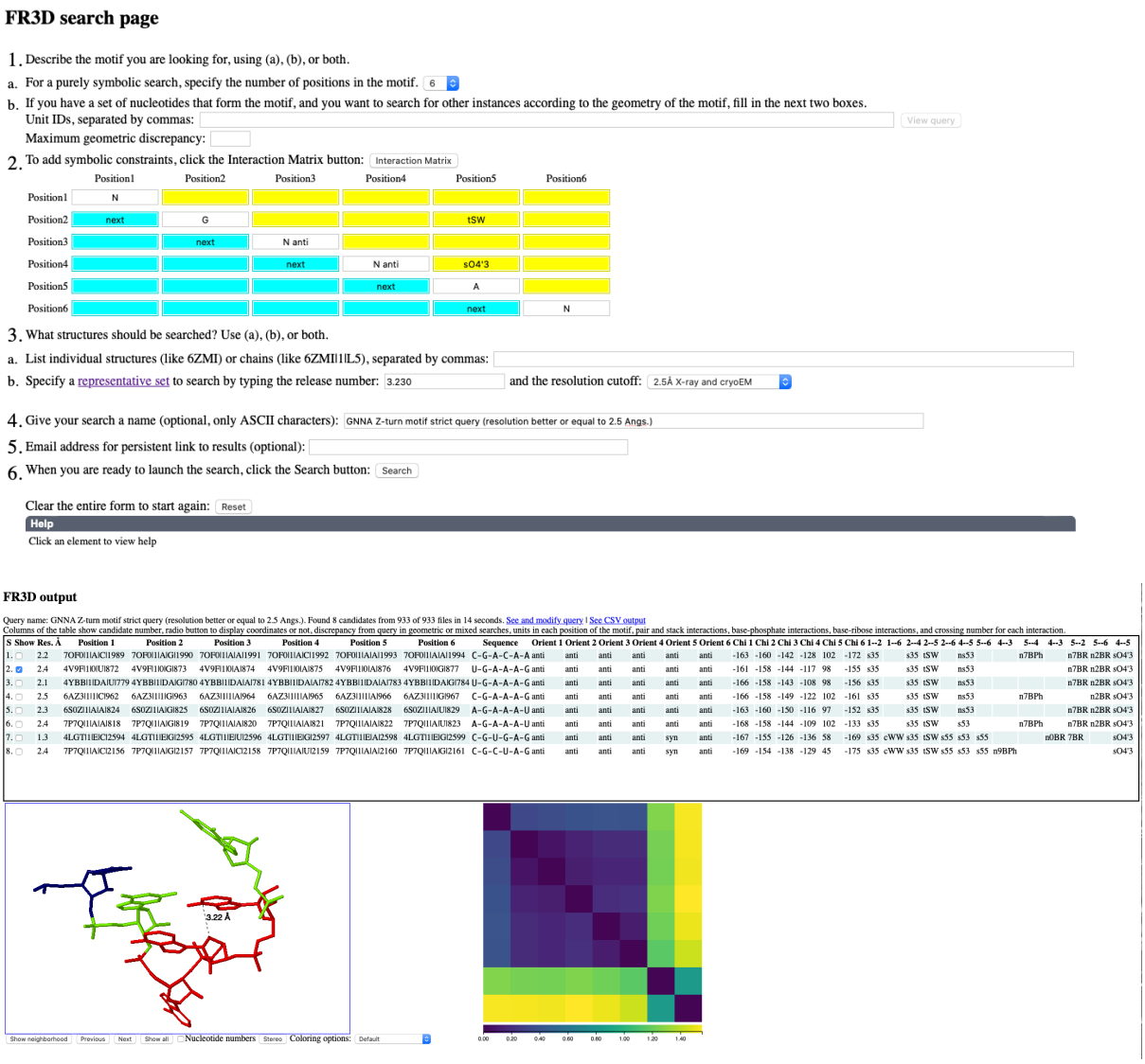

Figure S4: WebFR3D finds GNNA Z-turns. (Top) To search r(GNNA) Z-turns, we change the base identity in positions 2 and 5 but otherwise impose the same constraints as in **Figure S3**. (Bottom) The search retrieves several instances of r(GNNA) sequences that adopt a Z-turn conformation with an ApA Z-step. Besides, it finds a r(GACA) and a r(GUGA) Z-turn. A sO4'3 contact to the (A)C2 atom is visualized by a dashed line. Use the following link to retrieve the results of this search:
http://rna.bgsu.edu/webfr3d/Results/6280cd2b79770/6280cd2b79770.html

(last accessed the 6th of July 2022)

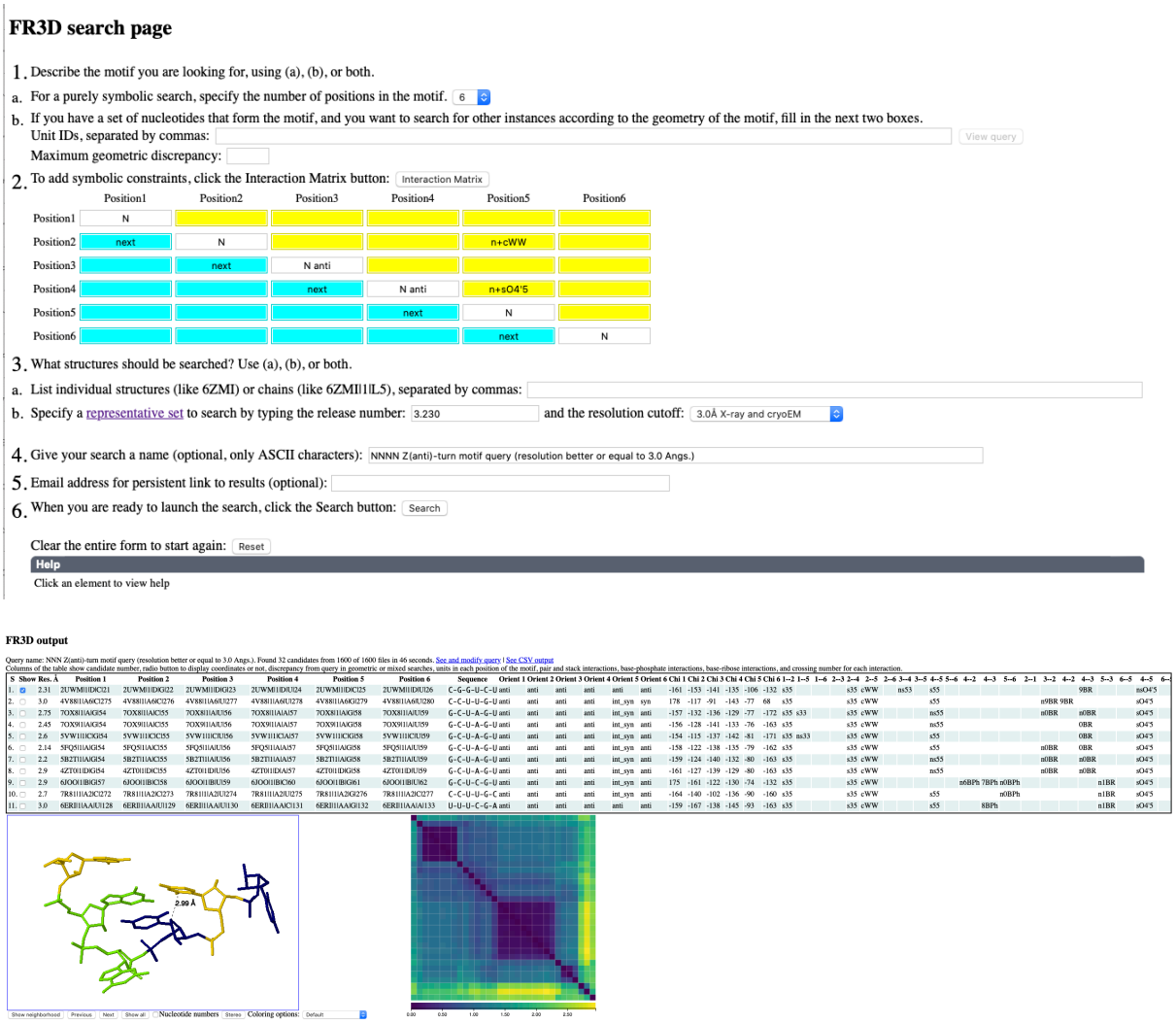

Figure S5: WebFR3D finds Z_{anti}-turns that fold similarly to Z-turns. (Top) To search for r(NNNN) Z_{anti}-turns, we follow the general model of **Figure S3**. However, we need to impose a nt2…nt5 cWW constraint and we need to change the nt4…nt5 sO4'3 to sO4'5 given the *anti* versus *syn* orientation of the 5'-nt. We include near versions of these interactions to retrieve a wider range of instances. (Bottom) A search on structures with resolution ≤ 3.0 Å retrieves 32 instances of Z_{anti}-turns. The structure shown is a surprising r(GGUC) Z_{anti}-turn; the shortest contact distance of the O4' atom with the cytosine ring involves the C2 atom and is shown by a dashed line. To retrieve this contact, we used the "near" syntax, i.e., n+sO4'5 (see **Method section**). Use the following link to retrieve the results of this search:
http://rna.bgsu.edu/webfr3d/Results/628109954b02d/628109954b02d.html

(last accessed the 6th of July 2022)

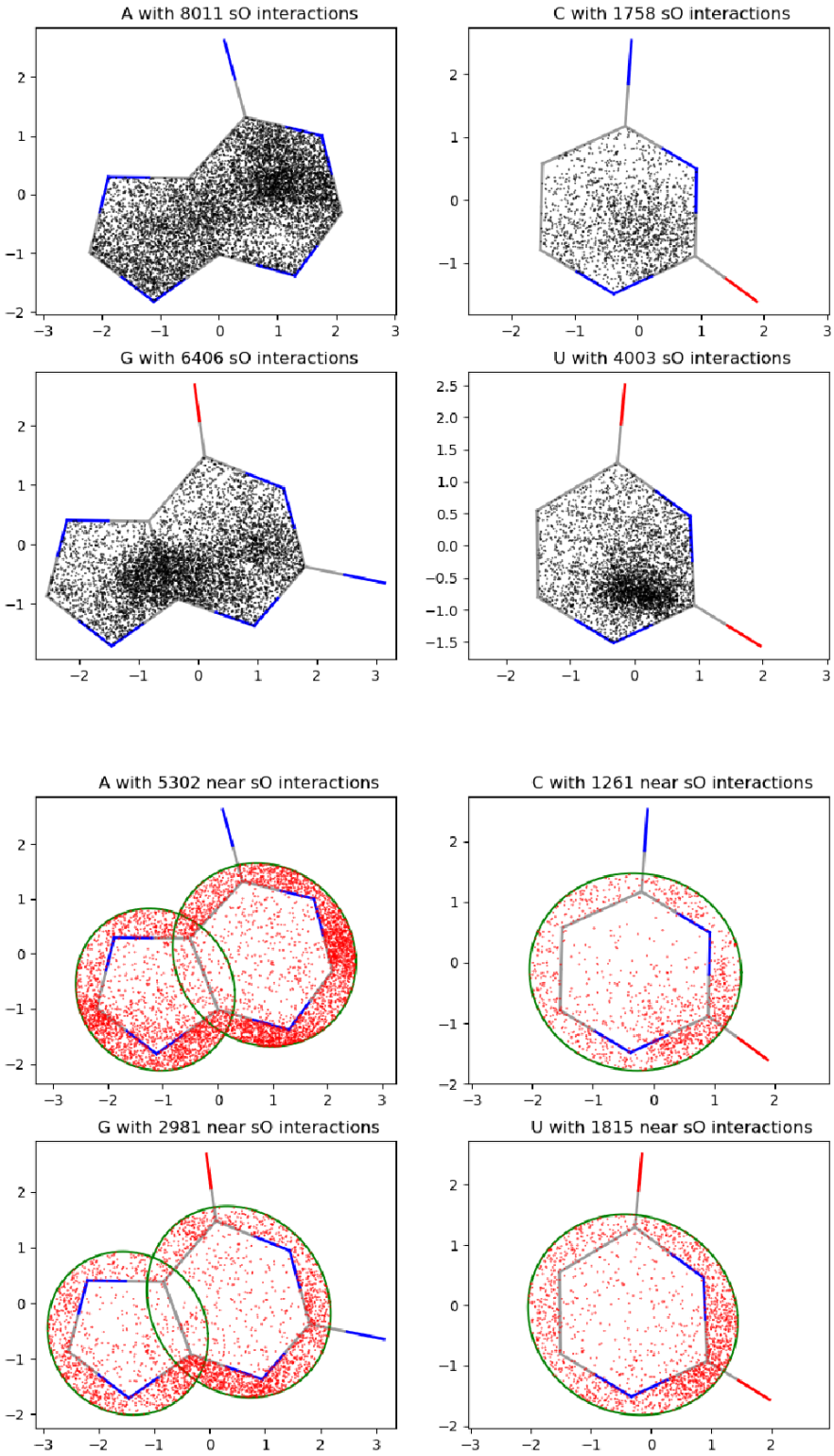

Figure S6: WebFR3D hits as visualized over the four nucleobases with a sO search at 3.0 Å resolution. (Top panels) Oxygen locations projected onto the plane of the base, obtained with a sO search as shown in **Figure S1** for all four nucleobases from an ensemble of structures with resolution ≤ 3.0 Å. (Bottom panels) These panels show the projected oxygen locations in "near" hits obtained with a nsO constraint and explain the working of the "near" keyword.

References

1. Watson, Z. L.; Ward, F. R.; Meheust, R.; Ad, O.; Schepartz, A.; Banfield, J. F.; Cate, J. H., Structure of the bacterial ribosome at 2 Å resolution. *Elife* **2020,** 9.
2. D'Ascenzo, L.; Leonarski, F.; Vicens, Q.; Auffinger, P., Revisiting GNRA and UNCG folds: U-turns versus Z-turns in RNA hairpin loops. *RNA* **2017,** 23, 259-269.