**SUPPLEMENTARY DATA**

***for***

**Secondary Structure Motifs Made Searchable to Facilitate the Functional Peptide Design**

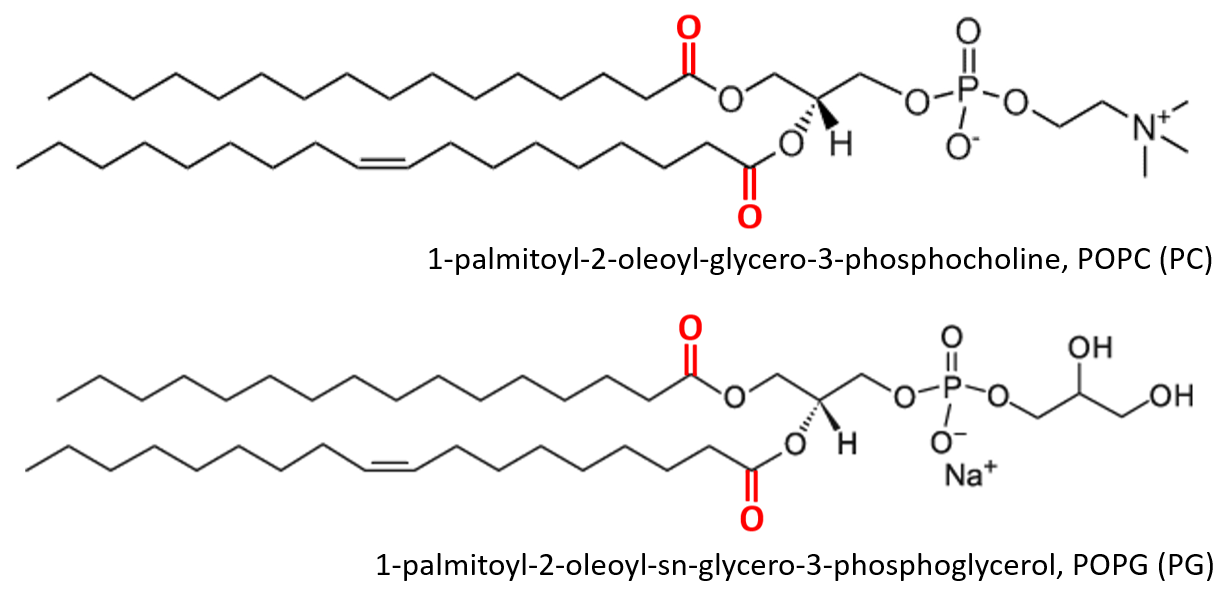
Cheng-Yu Tsai1,2,¶, Emmanuel O Salawu1,3,¶, Hongchun Li1,4,¶, Guan-Yu Lin5,¶, Ting-Yu Kuo5,¶, Liyin Voon1,¶, Adarsh Sharma1, Kai-Di Hu1, Yi-Yun Cheng6, Sobha Sahoo1, Lutimba Stuart1, Chih-Wei Chen1, Yuan-Yu Chang1, Yu-Lin Lu1, Ximai Ke1, Chen-Chi Wu2, Chung-Yu Lan5,7\*, Hua-Wen Fu5,7\*, Lee-Wei Yang1,3,7,8,\*

*1Institute of Bioinformatics and Structural Biology, National Tsing-Hua University, Hsinchu 30013, Taiwan; 2Graduate Institute of Medical Genomics and Proteomics, National Taiwan University College of Medicine, Taipei 10055, Taiwan; 3Bioinformatics Program, Institute of Information Sciences, Academia Sinica, Taipei 11529, Taiwan;* *4Department of Computational and Systems Biology, School of Medicine, University of Pittsburgh, Pittsburgh, PA, 15213, USA; 5Institute of Molecular and Cellular Biology, National Tsing Hua University, Hsinchu 30013, Taiwan;* *6Praexisio Taiwan Inc. New Taipei 22180, Taiwan; 7Department of Life Science, National Tsing Hua University, Hsinchu 30013, Taiwan*; *8Physics Division, National Center for Theoretical Sciences, Hsinchu 30013, Taiwan.*

¶ Authors share equal contributions

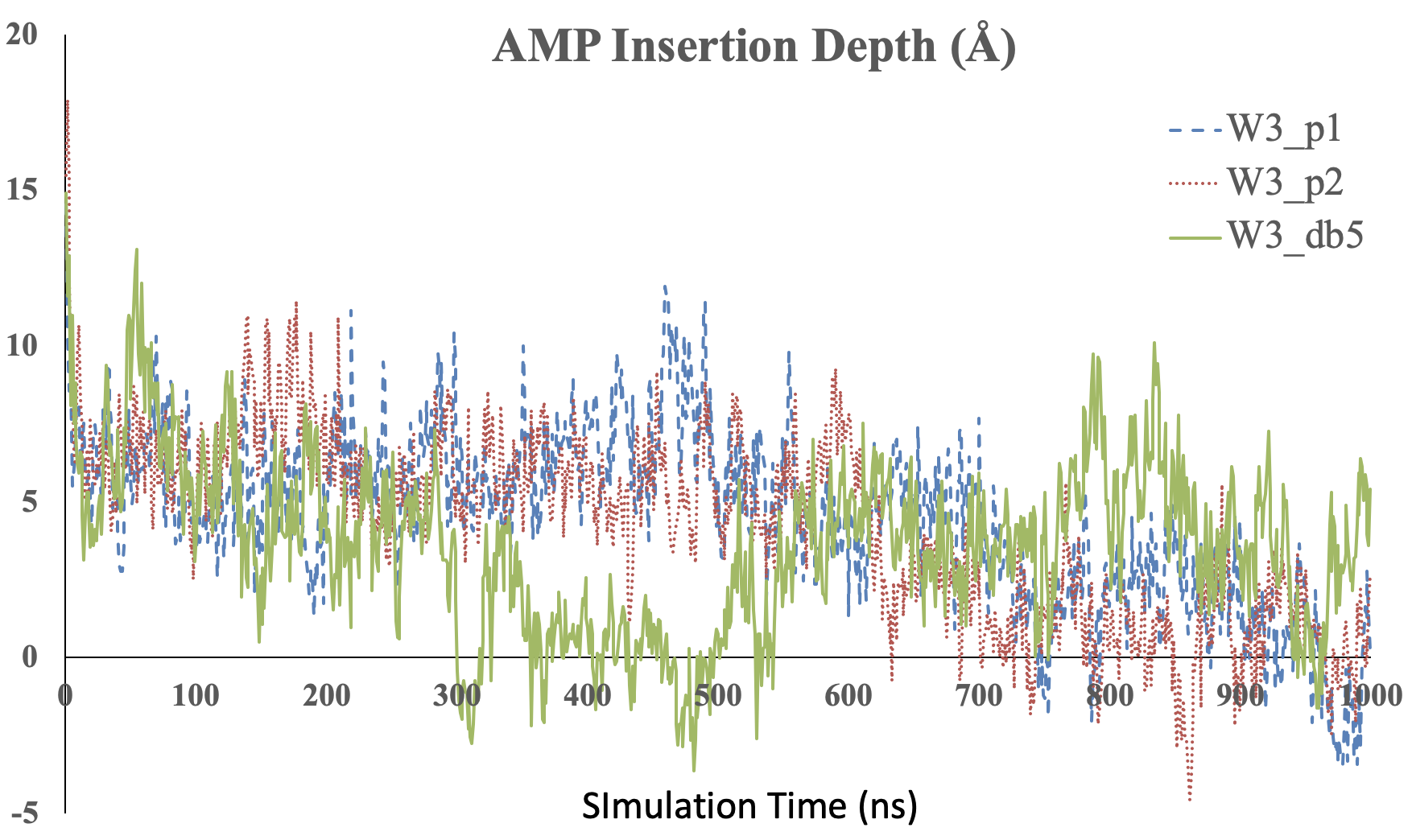
**Supplementary Results**

*Chemical structures of POPC and POPG lipids*



**Figure. S1** Chemical structure of POPC (PC) and POPG (PG). All images modified from Avanti® product page (Product ID: 840457 and 850457). The carbonyl oxygen in the lipids are highlighted in red, and these oxygen atoms are represented as red points in Fig. 3A. The atoms of the inserted AMPs below these oxygen atoms are considered to interact with the aliphatic tails of the lipids.

*Natural insertion of three AMPs into bacteria membrane*



**Figure S2.** The relative position of the center of mass (COM) of AMPs to the COM of phosphor atoms in the upper leaflet consisting of 40 POPC/POPG (3:1) lipid molecules, mimicking bacteria membrane, is plotted against the simulation time. All the MD simulations were conducted for 1 microsecond (us) at body temperature and normal pressure (1 atm) using OpenMM20 with CHARMM36 forcefield21,22.

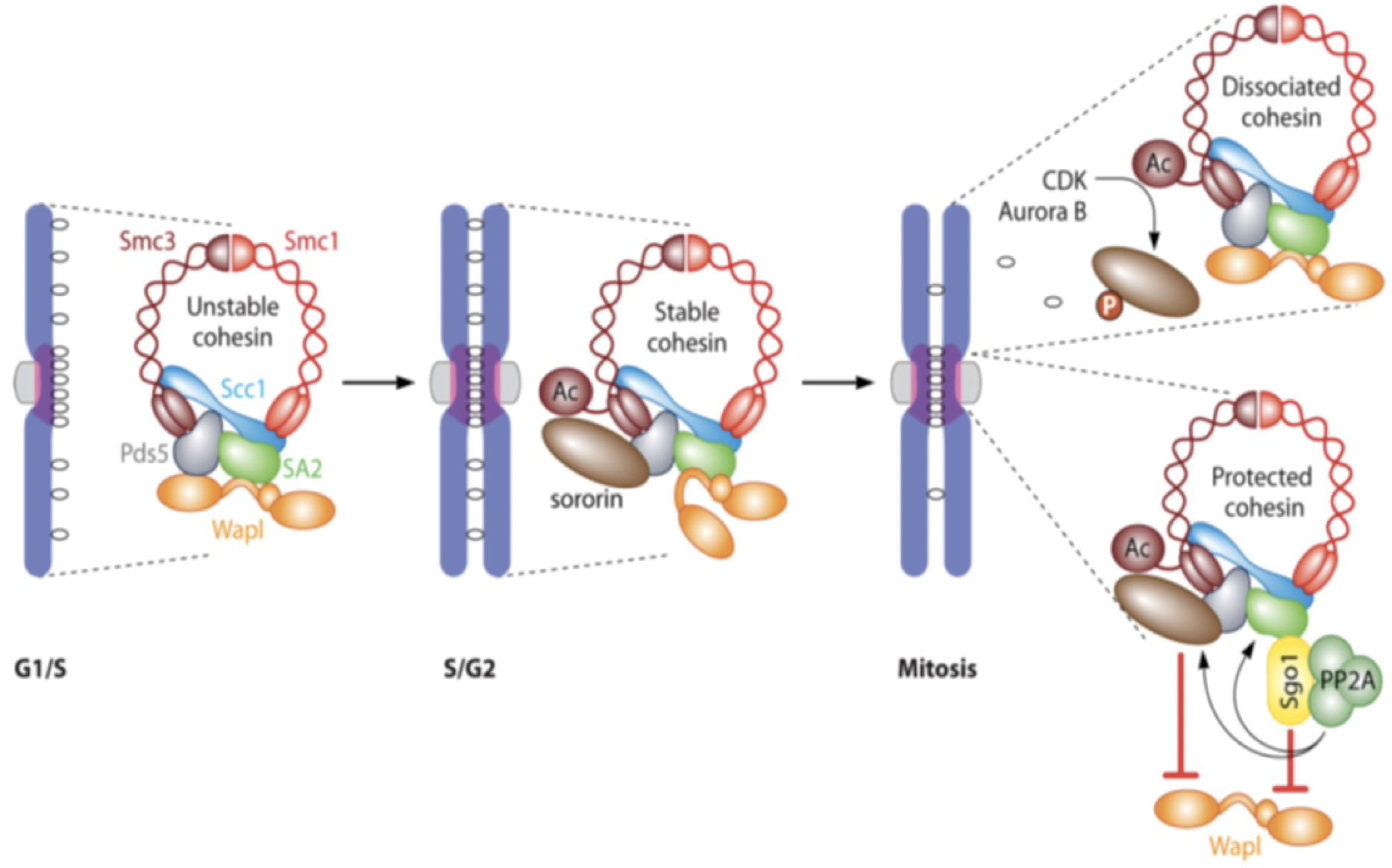
[**Supplementary Movie S1**](https://drive.google.com/file/d/1Ihm4OtY3u3QVdjhGY70x7BWHyTiMR7Q1/view?usp=sharing)

All the three AMPs were first created as fully extended AMPs and they were then simulated in an NPT ensemble for 50 ns in the absence of lipids. The resulting snapshots of the AMPs were clustered into 5 groups based on their structural similarity using the “clustering” plug-in of the VMD software31. The representative conformation from the biggest cluster in each of the three AMPs was selected as the initial conformation to be simulated from time zero in an NPT ensemble in the presence of the already equilibrated PC/PG (3:1) membrane. At time zero, as seen in the video, all the peptides have been folded into a helical shape in the CHARMM36 forcefield21,22. Within 1 microsecond, the tryptophan residues in W3\_p1 had all three Trp and W3\_p2 had two Trp (W3 and W6) inserted into the membrane (especially after 800 ns), while only one Trp (W3) in W3\_db5 had been inserted into the membrane.

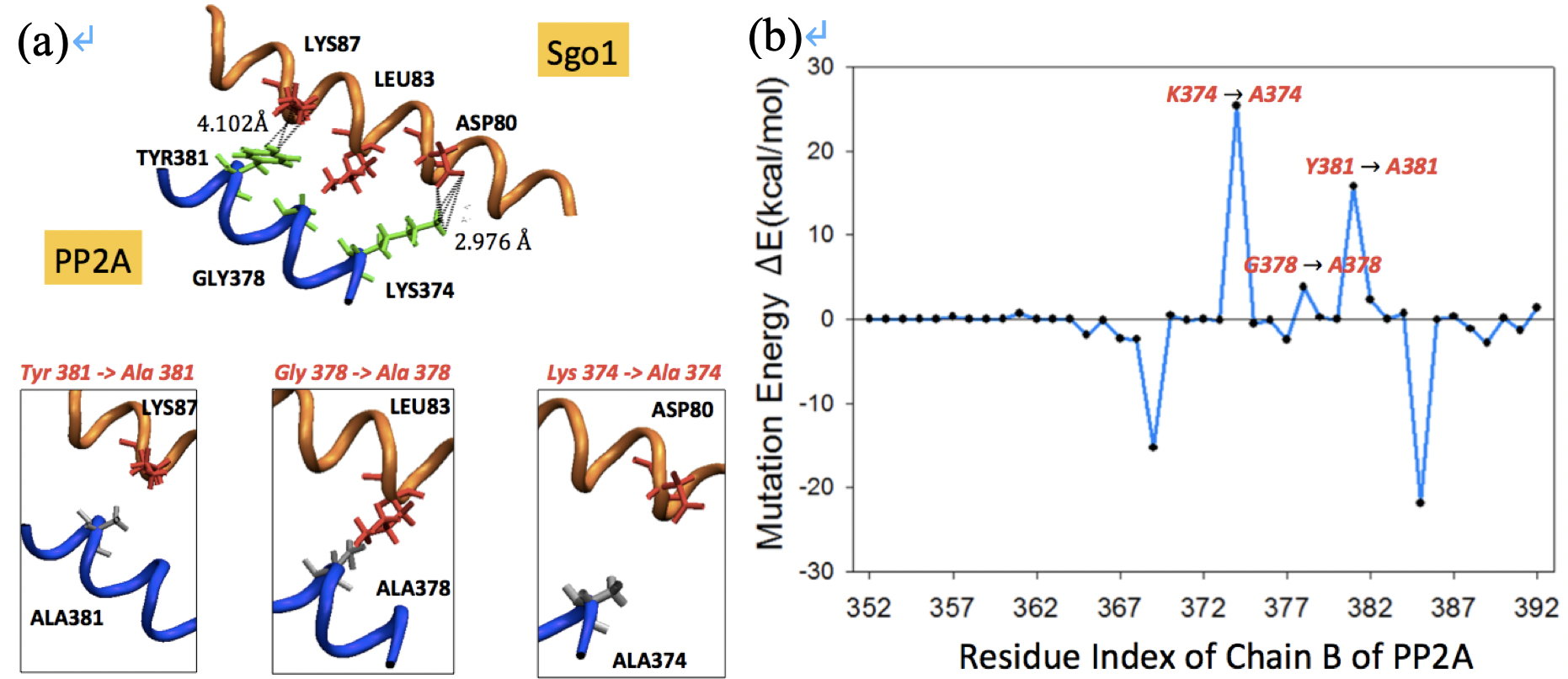
*Design of peptide inhibitor against Sgo1-PP2A interaction*

Below we demonstrate how TP-DB can be used to design helical blockers to prevent disease-related protein-protein interaction.

In late-stage hepatocellular carcinoma (HCC), patients rely only on targeted therapy and chemotherapy, which were proved generally ineffective. Novel therapeutic targets for treating HCC are urgently needed. Researchers previously found that protein Shugoshin-1 (Sgo1), protecting the stability of centromeric cohesin at centromere and ensuring proper chromosome separation55, was up-regulated in HCC. Notably, hepatoma cells were found more sensitive to Sgo1 deficiency than normal hepatocytes56. These results could suggest that Sgo1 can be a potential therapeutic target for HCC. It has been known that phosphorylated-Sgo1 recruits protein phosphatase 2A (PP2A) to centromere during early mitosis57 and the Sgo1-PP2A complex sequestered SA2 subunit of cohesin from Plk1-mediated phosphorylation to maintain its stability58. Additionally, Sgo1-PP2A complex also maintains Sororin in a hypo- phosphorylated state and keeps the interaction of Sororin-Pds5 to counteract WAPL57. These regulations ensure centromeric cohesin tethering sister chromatids together, until the onset of anaphase (Fig S3). As a result, an inhibitor (possibly a peptide) that blocks the Sgo1-PP2A interaction could potentially suppress the rapidly growing hepatoma cells, susceptible to improper Sgo1-PP2A association, in mitosis.

**Figure S3** The overview of Sgo1-PP2A protecting centromeric cohesin. During mitosis, phosphorylation of SA2 and Sororin cause cohesin to dissociate from chromosome arm. Cdk1-mediated phospho-Sgo1 recruits PP2A at centromere to counteract phosphorylation of SA2 and Sororin and maintains centromeric cohesin to tether sister chromatids together. Image is adopted from Figure 3 in Marston’s review article59.

We noted that the complex of Sgo1-PP2A has been solved by x-ray crystallography (PDB ID: 3FGA) and the two molecules bind with each other through an helix-helix interface (one helix from each molecule)11. To design anticancer peptides as blocker against Sgo1-PP2A interaction, the helical stretch in PP2A (residues from 352-392 in chain B of PDB 3FGA), known to form the helix-helix interaction with another helix in Sgo1 (residues from 75 to 92 in chain D of PDB 3FGA) (Fig S4), is chosen as the candidate target for further investigation. Evidenced from the x-ray structure, the helix-helix interaction is visibly apparent to be the most important essential interacting elements at the protein-protein interface. As shown in Figure S3a, K374, G378 and Y381 in the helical stretch “KTIHGLIY” could be the anchoring residues on PP2A to mediate the Sgo1-PP2A interaction, which is supported by our calculation of energetics using *in silico* alanine scanning (Fig S4b).

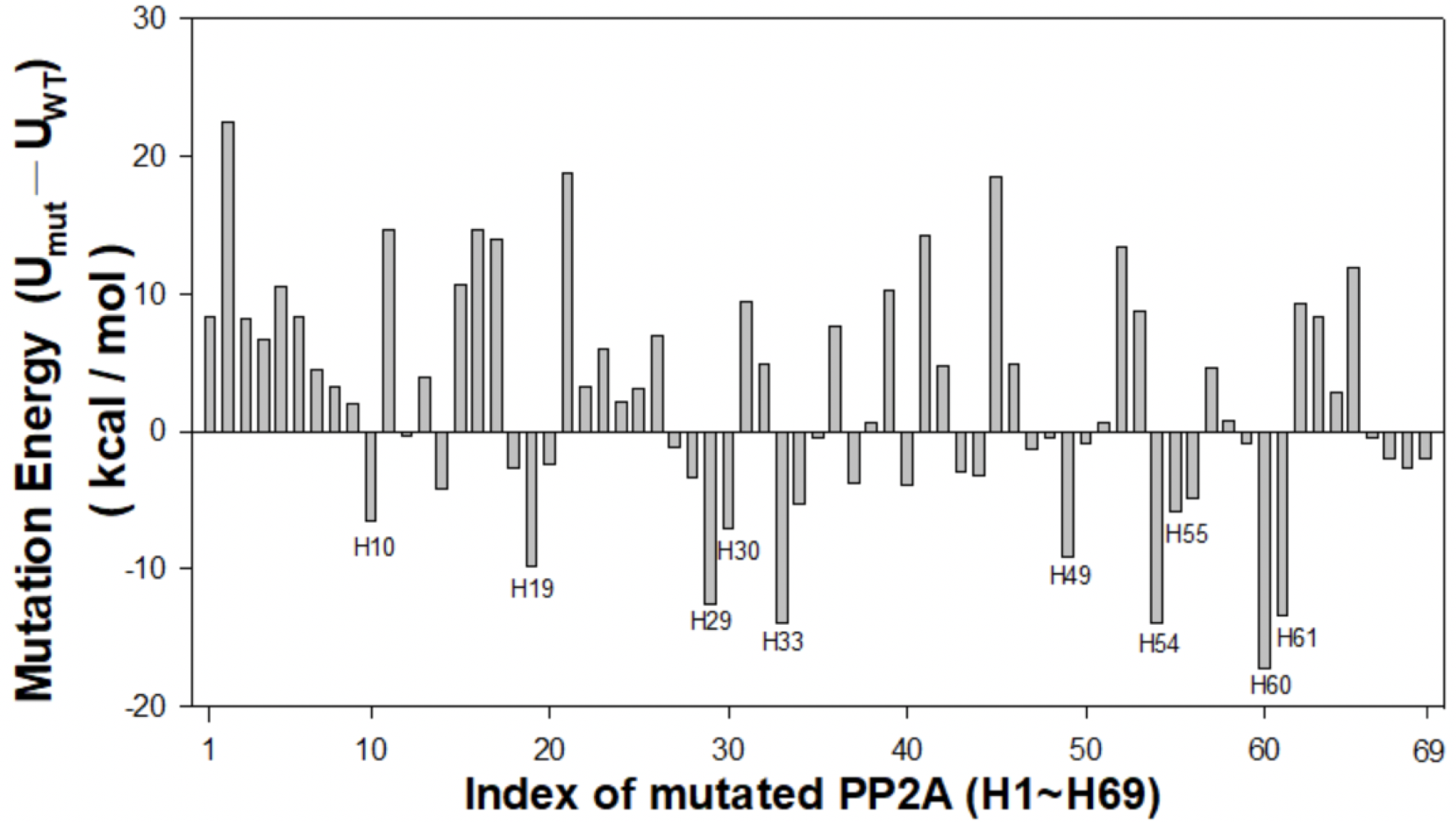


**Figure S4** Protein-protein interaction (PPI) between Sgo1 and PP2A through helix-helix interactions. Panel **(a, upper)** shows only the PPI interface between Sgo1 and PP2A, while the rest of proteins (3FGA; Xu et al., 2009) are hidden for clarity. Sgo1-PP2A interaction is mediated by two helices, one from Sgo1 (PDB: 3FGA; V75KEAQDIILQLRKECYYL92 in chain D) and the other from PP2A (PDB: 3FGA; K369THMNKTIHGLIYNALK385 in chain B) . Computational alanine scanning for every residue in the PP2A helical fragment (from residue 352 to 392 of chain B) is carried out, and in panel (**b**) energetic contribution for each residue mutated into alanine is assessed (from residue 369 to 385 of chain B) where a large positive value (e.g. for K374A, G378A and Y381A) indicates the importance of the residues in binding, and a large negative value suggests mutation into alanine favors the binding. The non-bonded energy is evaluated by NAMD package (Phillips et al., 2005) using CHARMM 36 forcefield (Huang and MacKerell, 2013). Here in panel **(b)**, the mutation energy for the n-th residue, ∆En, is defined as ∆En = Yn - X0 where X0 is the non-bonded energy of original complex between Sgo1 and PP2A (for the 352-392 fragment), and Yn is the energy of the complex with the *n*-th residue in PP2A being mutated into alanine, after a short energy minimization.

To block the protein-protein interaction (PPI) between Sgo1 and PP2A, one could potentially modify part of the most essential interacting element in PP2A into a peptide that interacts with Sgo1 better than its template. As shown in Figure S4b, computational alanine scanning helps detecting a stretch from 370 to 384 in PP2A (chain B) dominantly contributing the affinity in the PPI of Sgo1-PP2A complex, especially from K374, G378 and Y381 (i.e. aforementioned essential interacting element in PP2A). We then used this helical stretch, or the extended one with its flanking region from residue 364 to 388, including two energetically favored point mutations, K369A and K385A, as the starting template.

To find the helical peptides that bind the interface helix of Sgo1 stronger than PP2A does, we search the TP-DB for PP2A-like helical peptides using the queries summarized in Table S1.

As detailed in Supporting Methods, we first used the pattern “**K\*\*\*G\*\*Y**” to search TP-DB and discover 69 unique PP2A-like helical peptides (Table S1 and S2); we then carried out molecular dynamics simulations (MD) to assess how these peptides interact with Sgo1 at the helix-helix interface. Helices H60, H54, H61, H33 and H29 (**Fig S5**) are energetically promising for further experimental confirmation. Additionally, we examined other sets of peptides that met the patterns consisting of three anchoring residues (“A\*\*\*K\*\*\*G” and “G\*\*Y\*\*\*A”) or consisting of at least four residues (“A\*\*\*K\*\*\*G\*\*Y”, “K\*\*\*G\*\*Y\*\*\*A”, “A\*\*\*K\*\*\*G\*\*Y\*\*\*A” and “A\*\*/\*\*\*K\*\*/\*\*\*G\*\*/\*\*\*Y\*\*/\*\*\*A”**)** in TP-DB. All the corresponding results obtained from the queries are illustrated in **Table S3** to **S8**. Among them only the “A\*\*\*K\*\*\*G\*\*Y\*\*\*A” pattern (according to Figure S3b) returns no result from TP-DB.

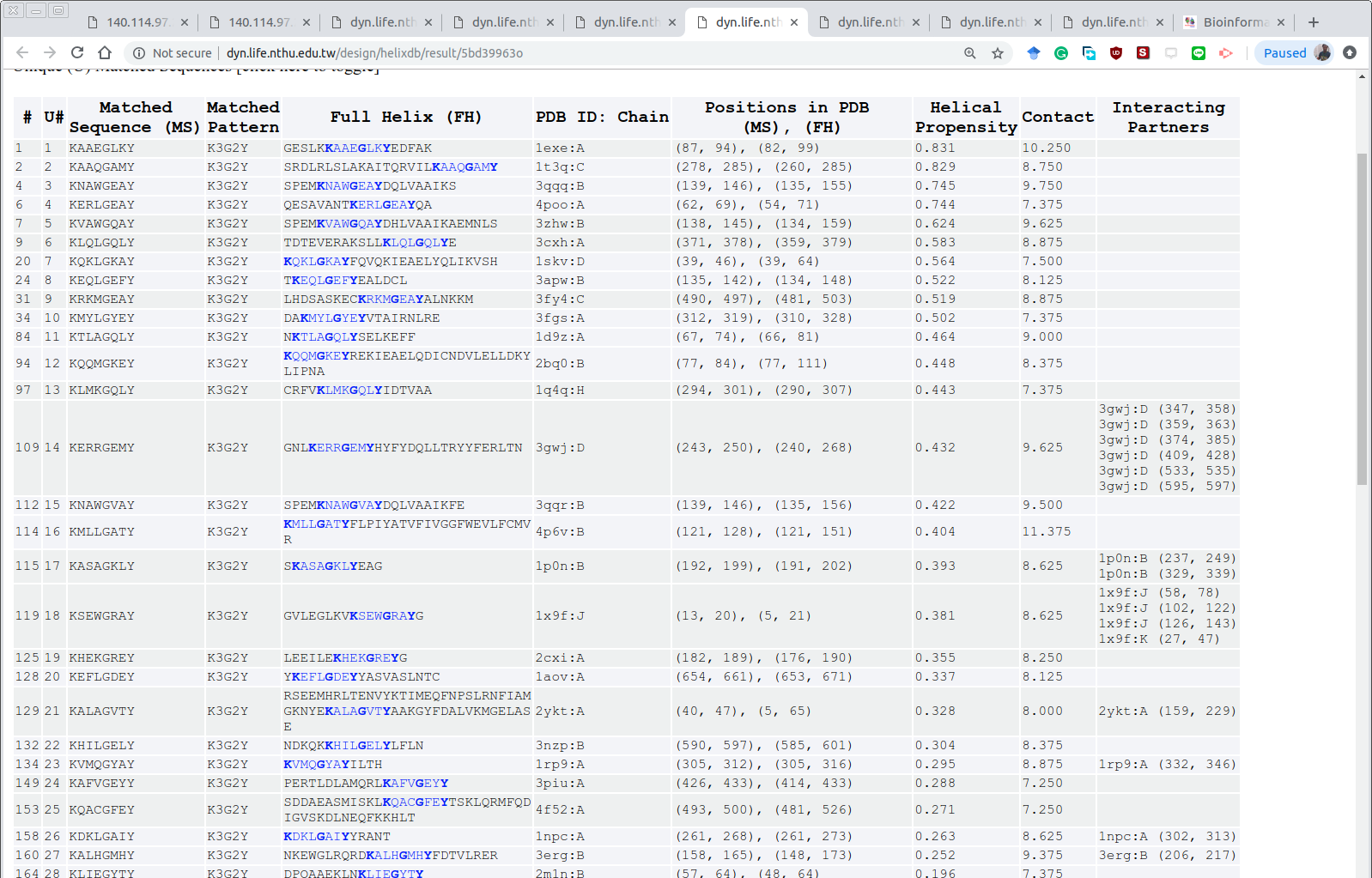


**Figure S5.** **Interactions between Sgo1 helix and PP2A-like peptides.** Umut is the interaction energy between targeted Sgo1 helix and each of the 69 PP2A-like peptides (from H1 to H69) that match the pattern “K3G2Y”. In detail, the energy is calculated by replacing the residue T375, I376, H377, L379 and I380 in the “**K374**TIH**G**LI**Y381**” stretch of PP2A with the corresponding residues in each of the 69 pattern-matched (K3G2Y) helical peptides returned from TP-DB. UWT is the interaction energy between Sgo1 helix and PP2A. Larger difference of interaction energies (Umut - UWT) suggest stronger binding of PP2A-like peptides with the Sgo1 helix than that between PP2A and the Sgo1 helix. Top ten values are further labeled in the graph.

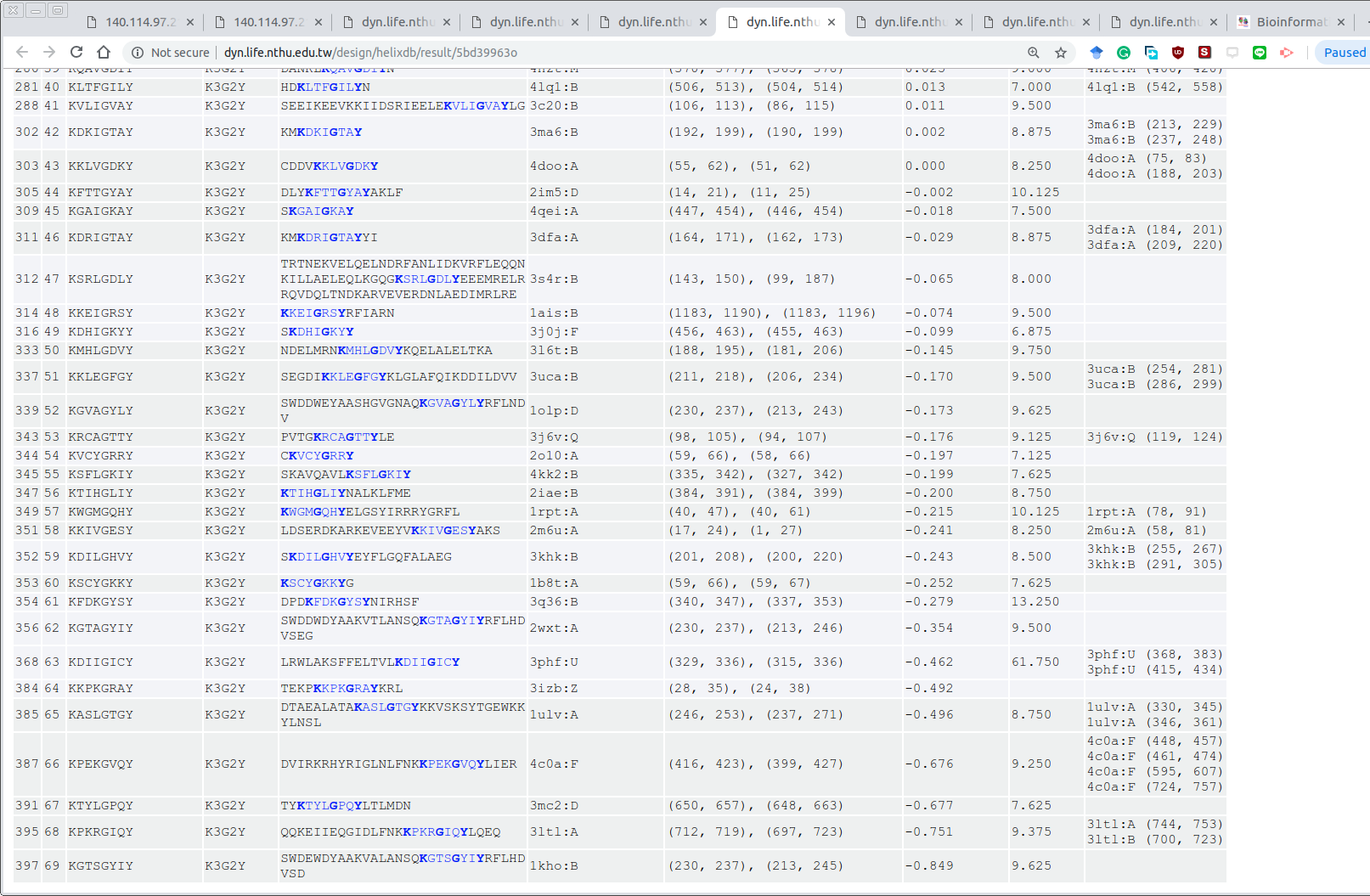
**Table S1**. PP2A-like patterns searched against the developed helical peptide database

|  |  |  |  |
| --- | --- | --- | --- |
| Pattern | Query | # of Peptides Found | # of Unique Peptides Found |
| K \*\*\* G \*\* Y | K 3 G 2 Y | 398 | 69 |
| A \*\*\* K \*\*\* G | A 3 K 3 G | 2981 | 391 |
| G \*\* Y \*\*\* A | G 2 Y 3 A | 2079 | 225 |
| A \*\*\* K \*\*\* G \*\* Y | A 3 K 3 G 2 Y | 15 | 3 |
| K \*\*\* G \*\* Y \*\*\* A | K 3 G 2 Y 3 A | 3 | 2 |
| A \*\*\* K \*\*\* G \*\* Y \*\*\* A | A 3 K 3 G 2 Y 3 A | 0 | 0 |
| A\*\*/\*\*\*K\*\*/\*\*\*G\*\*/\*\*\*Y\*\*/\*\*\*A | A 2,3 K 2,3 G 2,3 Y 2,3 A | 34 | 3 |

**Table S2. 69 unique helical sequences obtained from the TP-DB matching the pattern K\*\*\*G\*\*Y given the query of [K 3 G 2 Y]**

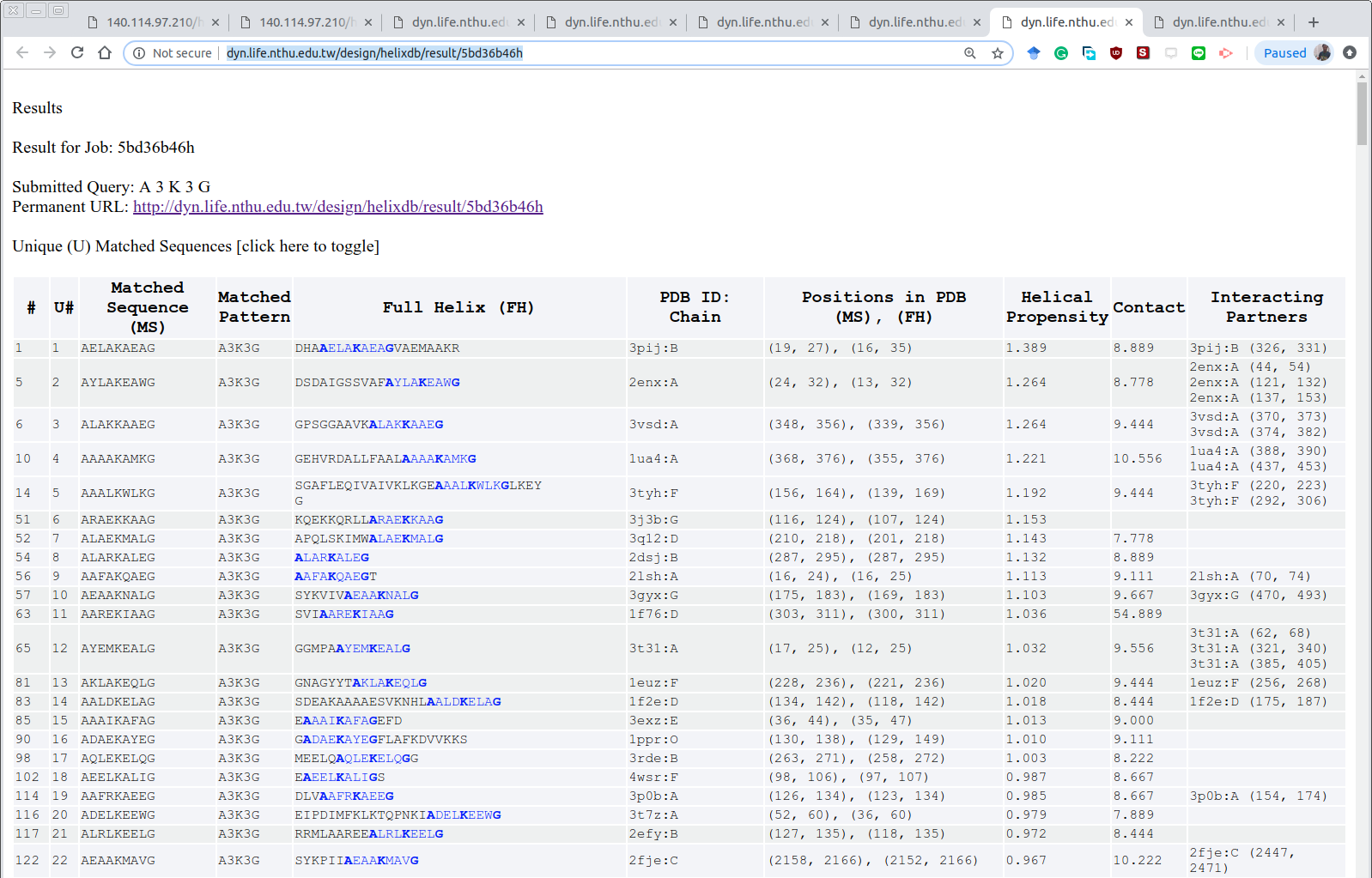


:

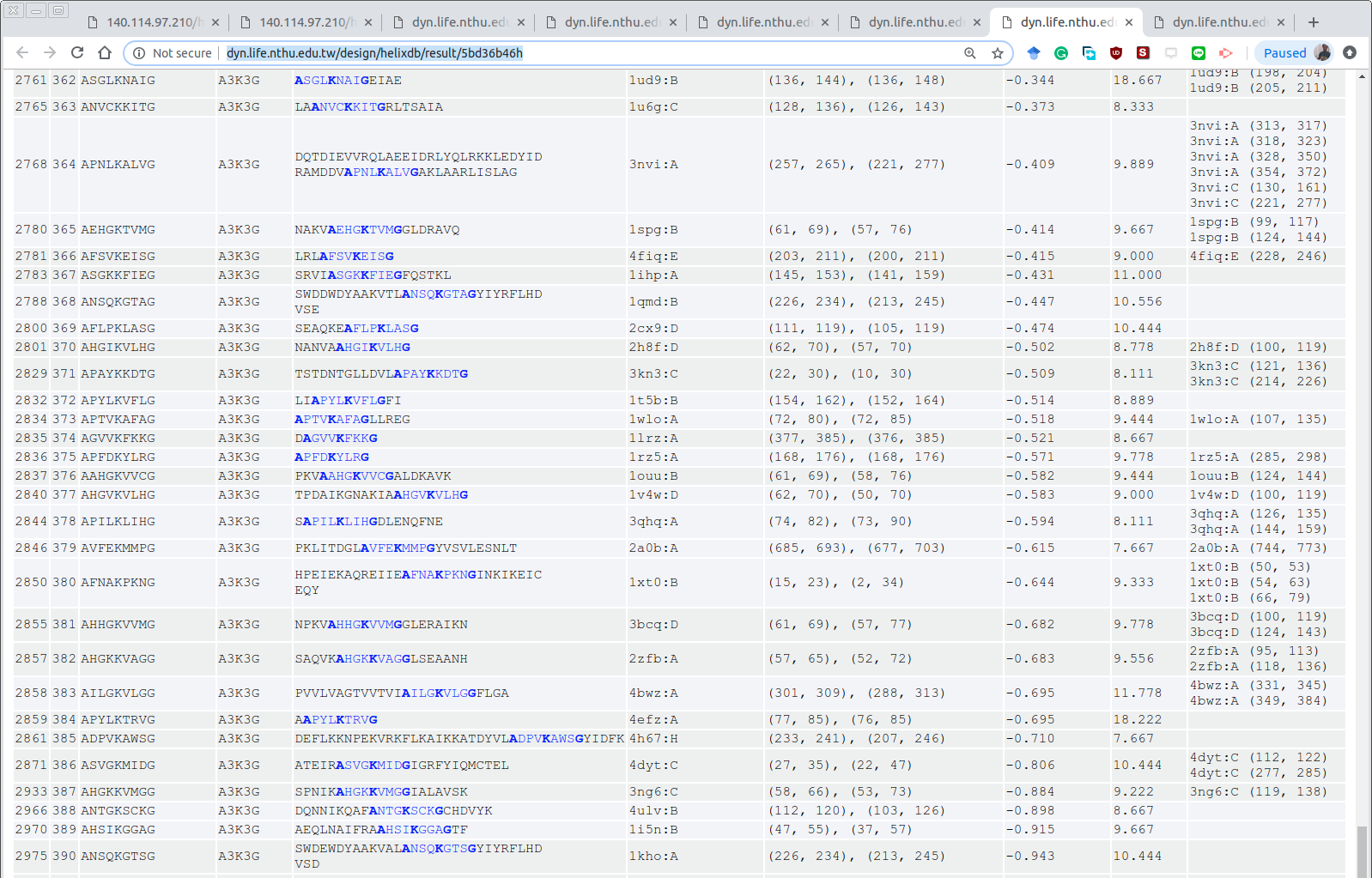


The result page is available online at <https://dyn.life.nthu.edu.tw/design/result?JobID=602bae01v>

**Table S3. 390 unique helical sequences obtained from the TP-DB for the pattern A\*\*\*K\*\*\*G given the query of [A 3 K 3 G]**

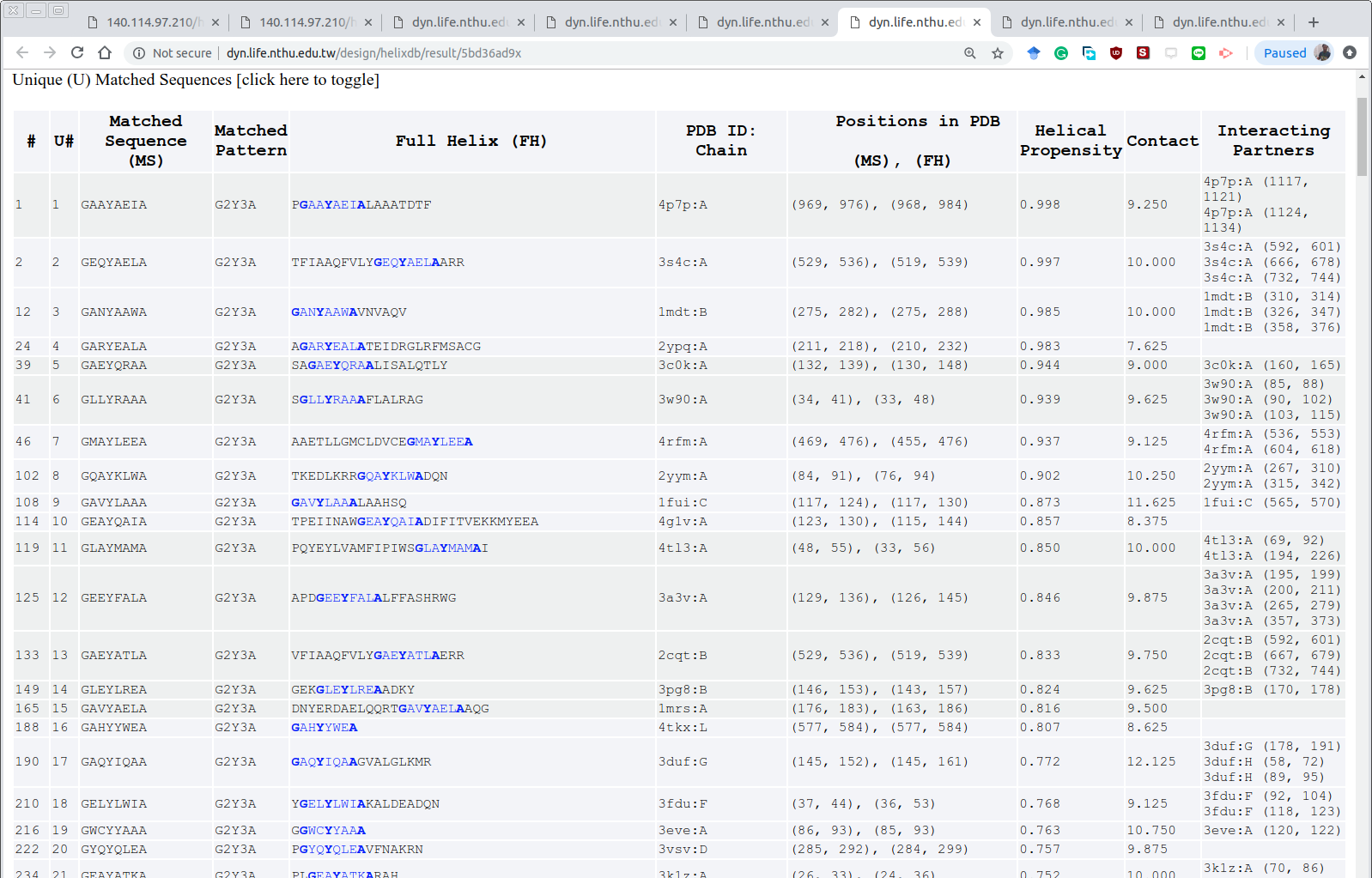


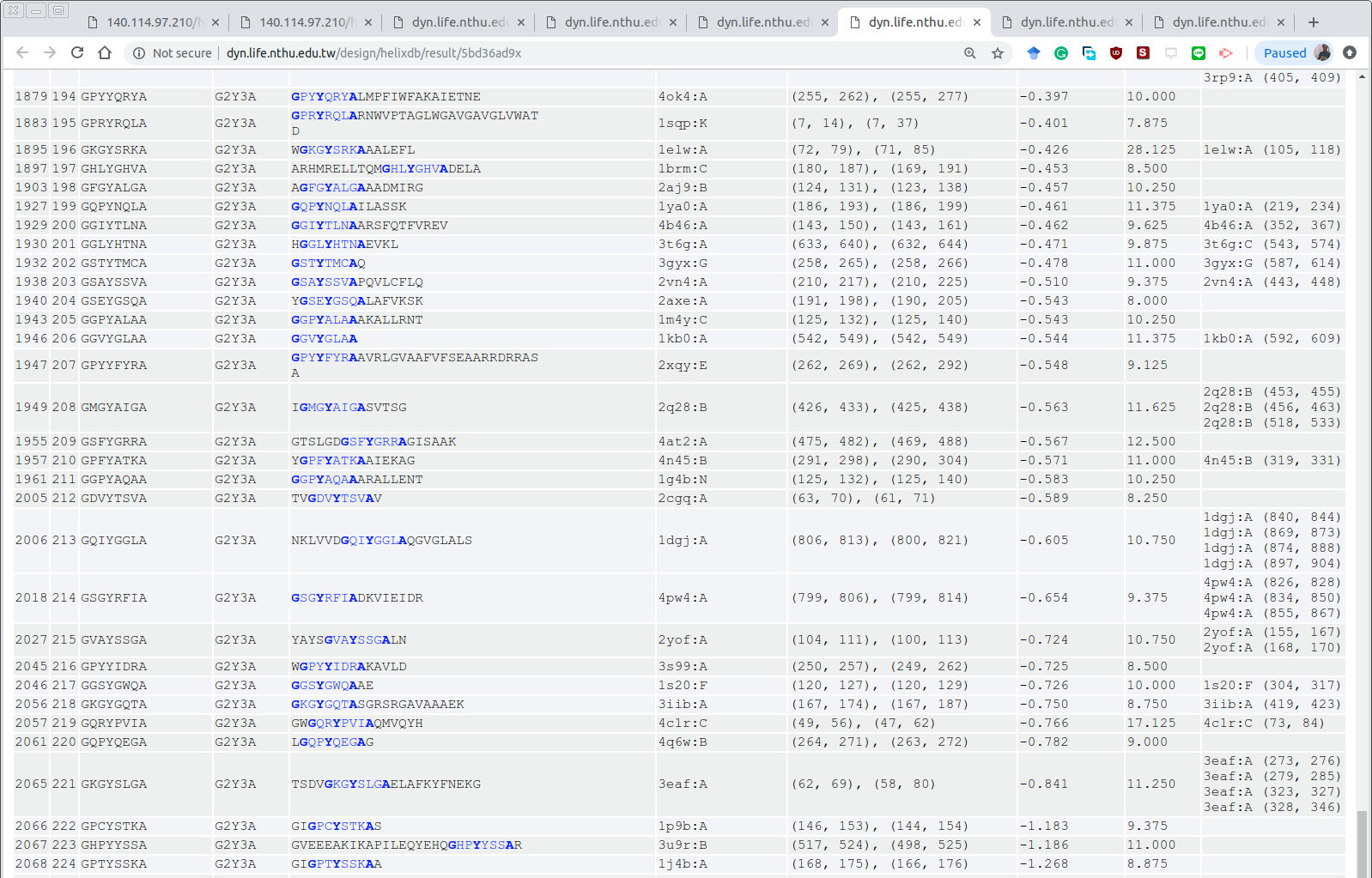
:



The result page is available online at <https://dyn.life.nthu.edu.tw/design/result?JobID=602bae5cu>

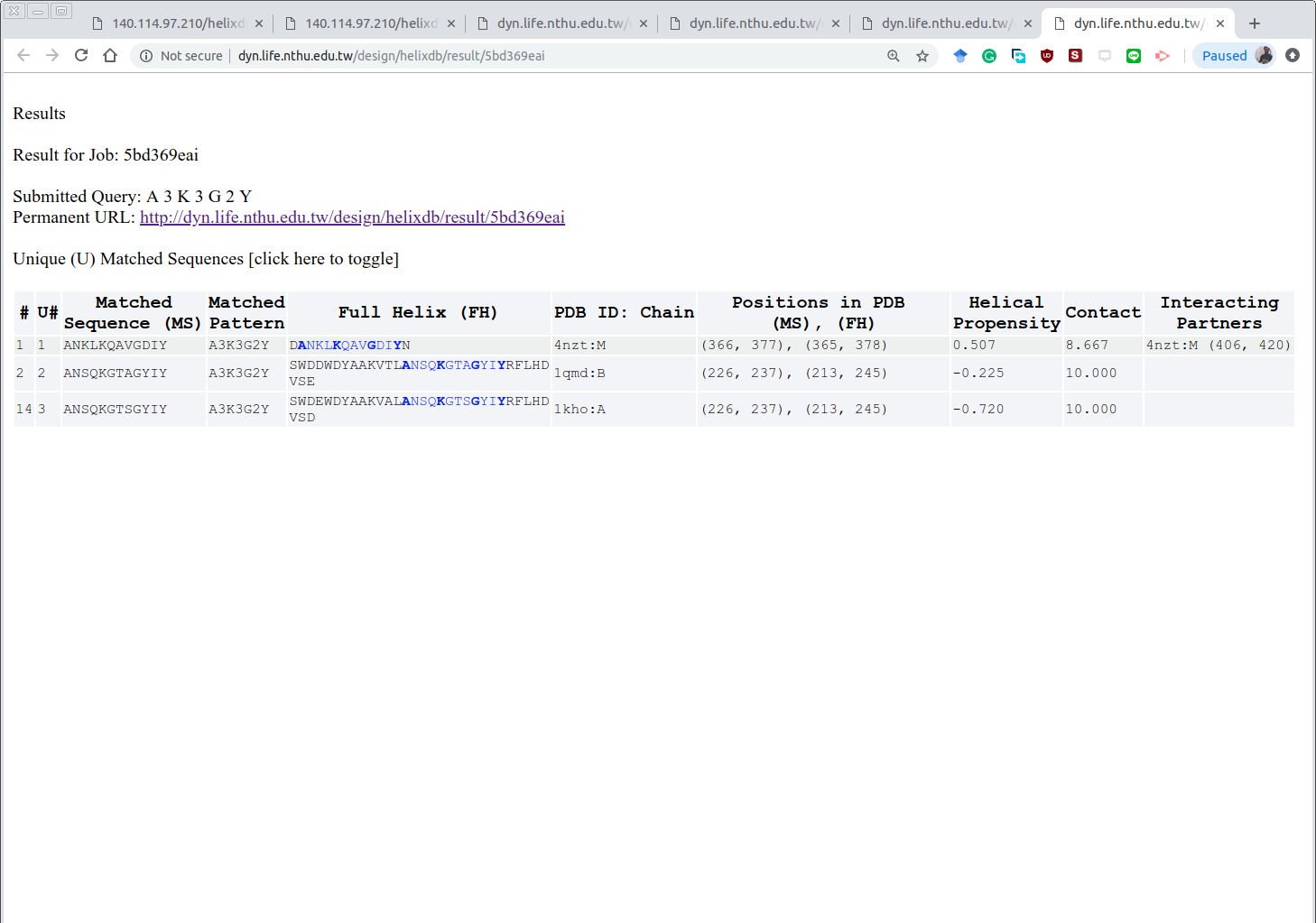
**Table S4. 224 unique helical sequences obtained from the TP-DB for the pattern G\*\*Y\*\*\*A given the query of [G 2 Y 3 A]**

:



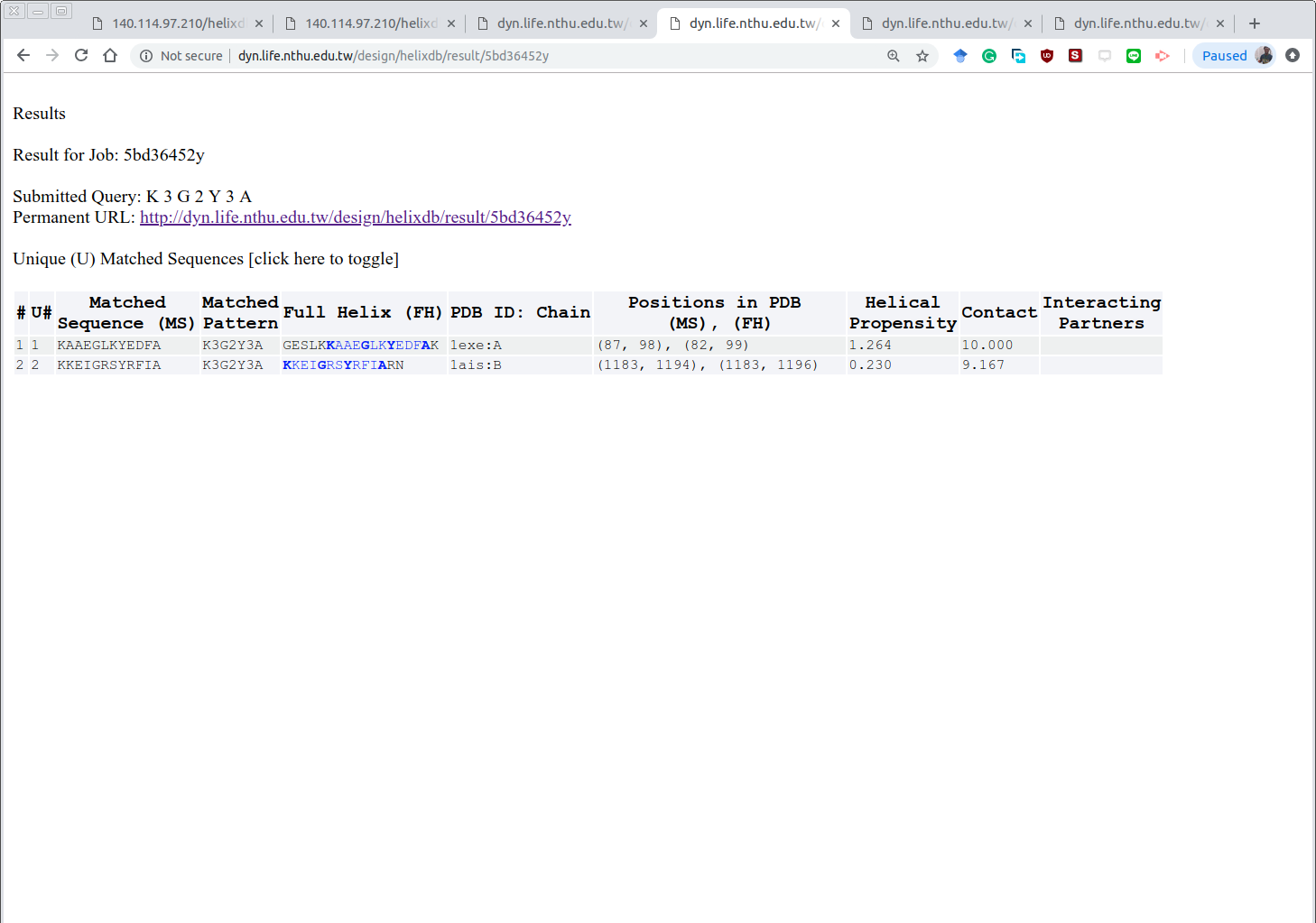
The result page is available online at <https://dyn.life.nthu.edu.tw/design/result?JobID=602bb16dr>

**Table S5. Results obtained from the TP-DB for pattern A\*\*\*K\*\*\*G\*\*Y with query [A 3 K 3 G 2 Y]**



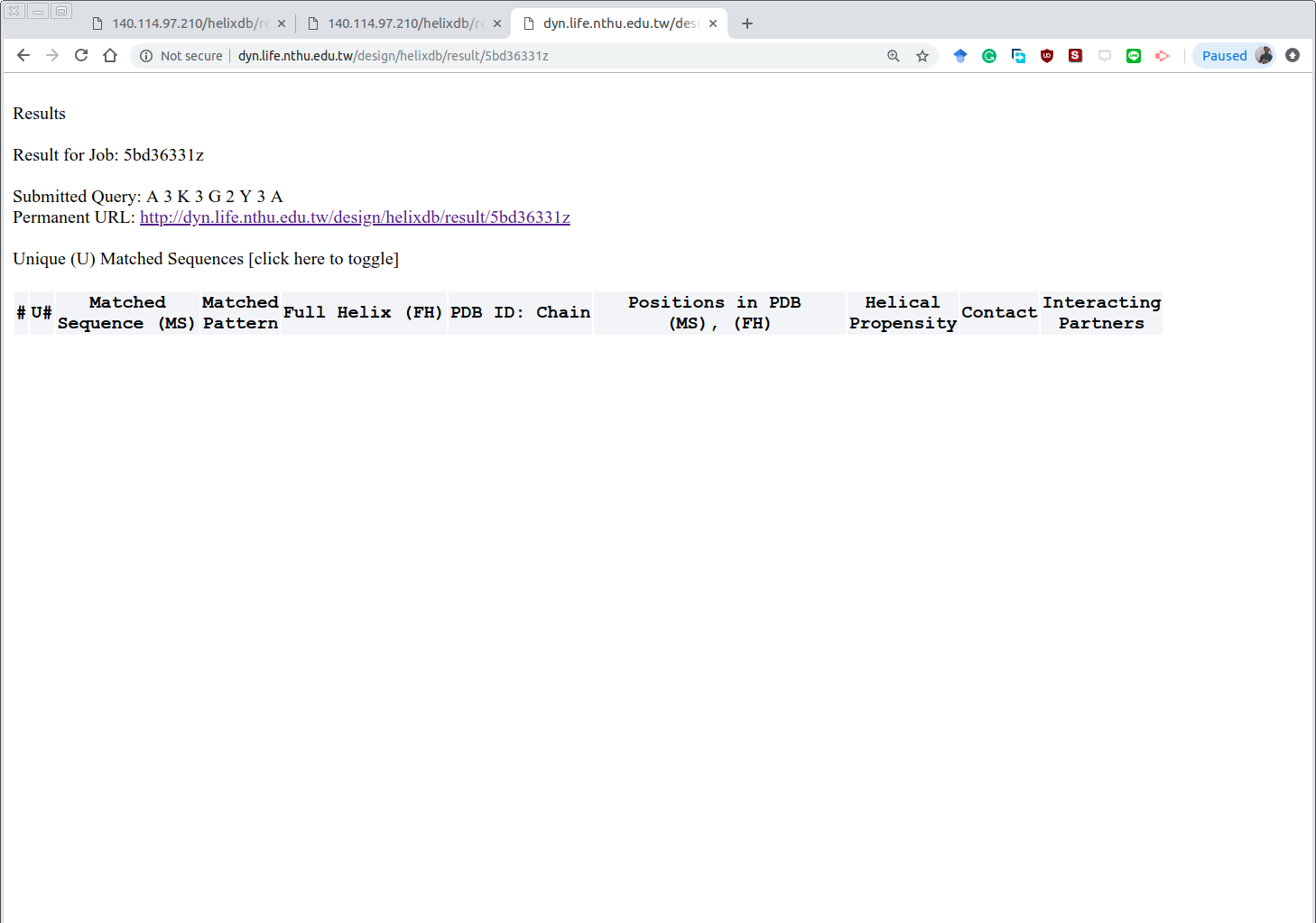
The result page is available online at <https://dyn.life.nthu.edu.tw/design/result?JobID=602bb1e0r>

**Table S6. Results obtained from the TP-DB for pattern K\*\*\*G\*\*Y\*\*\*A with query [K 3 G 2 Y 3 A]**



The result page is available online at <https://dyn.life.nthu.edu.tw/design/result?JobID=602bb2e9z>

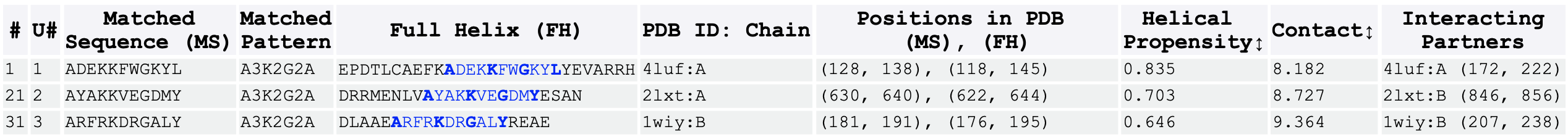
**Table S7. Results obtained from the TP-DB for pattern A\*\*\*K\*\*\*G\*\*Y\*\*\*A with query [A 3 K 3 G 2 Y 3 A]**



No peptide found

The result page is available online at <https://dyn.life.nthu.edu.tw/design/result?JobID=602bb312l>

**Table S8. Results obtained from the TP-DB for pattern** **A\*\*/\*\*\*K\*\*/\*\*\*G\*\*/\*\*\*Y\*\*/\*\*\*A with query [A 2,3 K 2,3 G 2,3 Y 2,3 A]**



The result page is available online at <https://dyn.life.nthu.edu.tw/design/result?JobID=602bb335c>

We selected the helical PP2A-like peptides found from queries [A 3 K 3 G 2 Y] (**Table S5**), [K 3 G 2 Y 3 A] **(Table** **S6**), [A 2,3 K 2,3 G 2,3 Y 2,3 A] (**Table S8**), as well as the wide-type PP2A helix as control, for further investigations through all-atom MD-based binding free energy change (∆G) evaluation by MM/PBSA (see Supporting Methods). The results (**Table S9**) showed that seven of the newly found helixes, with a binding ∆∆G < 0.0 kcal/mol, bound Sgo1 stronger than the wide-type PP2A did, suggesting a potential PPI blocking function for these peptides. Among them, the number one peptide “**A**YA**K**KVE**G**DM**Y**”, with a binding affinity ~10 kcal/mol stronger than the wild-type peptide, can be a promising peptide subject to further experimental validation by isothermal calorimetry (ITC) assays and/ or NMR HSQC spectra. Peptide design is an elaborated miniature of protein design, and we expect that herein proposed methodologies can be applied for the design of therapeutic α-helical peptides for other medicinal purposes.

**Table S9. Seven of the newly found helixes, with binding ∆∆G less than 0.0 kcal/mol, have the potentials of binding stronger to SGO1's helix than the control**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Candidate Helix  (Matched Pattern) | Helix Source, PDB ID: Chain | Helical Propensity Score | Contact  Number | Binding ∆∆G (kcal/mol) | In which Table above\* (U#) |
| **A**YA**K**KVE**G**DM**Y**  (**A 2,3 K 2,3 G 2,3 Y**) | 2lxt:A | 0.703 | 8.545 | -10.11 ± 0.43 | **Table S8** (2') |
| **K**KEI**G**RS**Y**RFI**A**  (**K 3 G 2 Y 3 A**) | 1ais:B | 0.230 | 9.167 | -9.20 ± 0.32 | **Table S6** (2) |
| **K**AAE**G**LK**Y**EDF**A**  (**K 3 G 2 Y 3 A**) | 1exe:A | 1.264 | 10.000 | -8.04 ± 0.29 | **Table S6** (1) |
| **A**YA**K**KVE**G**DM**Y**ES**A**  (**A 2,3 K 2,3 G 2,3 Y 2,3 A**) | 2lxt:A | 0.703 | 8.727 | -6.64 ± 0.45 | **Table S8** (2) |
| **A**NSQ**K**GTA**G**YI**Y**  (**A 3 K 3 G 2 Y**) | 1qmd:B | -0.225 | 10.000 | -5.63 ± 0.33 | **Table S5** (2) |
| **A**RFR**K**DR**G**AL**Y**RE**A**  (**A 2,3 K 2,3 G 2,3 Y 2,3 A**) | 1wiy:B | 0.646 | 9.364 | -1.96 ± 0.46 | **Table S8** (3) |
| **A**RFR**K**DR**G**AL**Y**  (**A 2,3 K 2,3 G 2,3 Y**) | 1wiy:B | 0.957 | 8.500 | -1.14 ± 0.38 | **Table S8** (3') |
| **K**TIH**G**LI**Y**  **Control, from PP2A** | 3fga:B | -0.200 | 9.250 | 0.00 |  |
| **A**NSQ**K**GTS**G**YI**Y**  (**A 3 K 3 G 2 Y**) | 1kho:A | -0.720 | 10.000 | 1.38 ± 0.39 | **Table S5** (3) |
| **A**DEK**K**FW**G**KYL**Y**  (**A 2,3 K 2,3 G 2,3 Y**) | 4luf:A | 0.957 | 8.500 | 4.31 ± 0.29 | **Table S8** (1') |
| **A**DEK**K**FW**G**KYL**Y**EV**A**  (**A 2,3 K 2,3 G 2,3 Y 2,3 A**) | 4luf:A | 0.835 | 8.182 | 4.66 ± 0.86 | **Table S8** (1) |
| **A**NKL**K**QAV**G**DI**Y**  (**A 3 K 3 G 2 Y**) | 4nzt:M | 0.507 | 8.667 | 5.11 ± 0.32 | **Table S5** (1) |

\*The earlier Tables that listed the matched sequences are denoted at the rightmost column. The number in the parenthesis is the serial number for the unique (U#) sequence in the corresponding Table of interest. A prime following a U#, such as 2', indicates that the sequence in that row of this table is a subsequence of the original sequence presented in the earlier Tables.

**Supplementary Methods**

*Production of recombinant Helicobactor pylori neutrophil-activating protein (HP-NAP) and maltose-binding protein (MBP)*

Recombinant *H.* *pylori* neutrophil-activating protein (HP-NAP) was expressed in *E. coli* BL21(DE3) cells harboring the expression plasmid pET42a-NAP and purified by either two consecutive gel-filtration chromatography as previously described60 or a small-scale DEAE Sephadex negative mode batch chromatography as previously described61. Maltose-binding protein (MBP) was prepared the same as the procedure for production of MBP fused with the polypeptide containing residues Arg77 to Glu116 of HP-NAP as described below except that *E. coli* BL21(DE3) cells harboring the pMALc2g expression vector was used for expression.

*Cloning of HP-NAPR77-E116 into a MBP fusion protein expression vector*

The plasmid DNA pET42a-NAP encoding a *napA* gene from *H. pylori* strain 26695 [GenBank:AE000543.1, Gene: HP0243] was prepared as previously described60. The DNA fragment coding for polypeptide containing residues Arg77 to Glu116 of HP-NAP (HP-NAPR77-E116), which contains the D-Y-K-x-x-[DE] motif, was amplified by PCR from the plasmid pET42a-NAP using the forward and reverse primers containing BamHI and HindIII site, respectively. The forward primer is 5’-ATAAGGATCCCGTGTTAAAGAAGAAACTAAAAC-3’ and the reversed primer is 5’-TTAATAAGCTTTAATTCTTTTTCAGCGGTGTTAGAG-3’. The PCR reaction was carried out with 10 ng plasmid DNA pET42a-NAP as a template and KAPA HiFi PCR Kit (Kapa Biosystems, Inc.) in a Mastercycler Gradient 5331 (Eppendorf, Germany). An initial denaturing phase of 95 °C for 5 min was followed by 39 cycles of 98 °C for 20 sec, 67 °C for 15 sec, and 72 °C for 15 sec. A final elongation phase of 72 °C for 2 min was also included. The amplified DNA fragments encoding HP-NAPR77-E116 were then cloned into pJET1.2/blunt vectors using the CloneJET PCR Cloning Kit (Thermo Fisher Scientific Inc.). The resulting plasmid was designated as pJET1.2/blunt-HP-NAPR77-E116. The insert was sequenced to confirm the correct DNA sequence. The correct insert was digested from pJET1.2/blunt-HP-NAPR77-E116 with BamHI and HindIII and then cloned into the pMALc2g expression vector62. The resulting plasmid was designated as pMALc2g-HP-NAPR77-E116.

*Production of MBP-tagged HP-NAPR77-E116*

*E. coli* BL21(DE3) cells harboring pMALc2g-HP-NAPR77-E116 were streaked on a lysogeny broth (LB) agar plate containing 100 μg/ml ampicillin and incubated at 37 °C for 16 hr. A single colony was picked and inoculated into 4 ml of LB containing 100 μg/ml ampicillin and the culture was incubated at 37 °C with shaking at 170 rpm for 16 hr. A volume of 2 ml of the overnight culture was inoculated into 200 ml LB containing 100 μg/ml ampicillin and the inoculated culture was incubated at 37 °C with shaking at 170 rpm for 2 hr until the OD600 reached 0.5. The expression of MBP-tagged HP-NAPR77-E116 was induced by the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.3 mM and the culture was incubated at 37 °C with shaking at 180 rpm for 3 h until the OD600 reached 1.7. Then, the cells were centrifuged at 6,000 x g at 4 °C for 15 minutes to remove the supernatant and the cell pellets were stored at -70 °C.

The cell pellet from a 200 ml culture of *E. coli* expressing recombinant MBP-tagged HP-NAPR77-E116 were re-suspended in 20 ml of ice-cold buffer containing 20 mM Tris-HCl, pH 7.4, 200 mM NaCl, 1 mM ethylenediaminetetraacetic acid (EDTA), and 1 mM dithiothreitol (DTT), plus 0.1% (v/v) protease inhibitor mixture (PI mix). The PI mix contained 0.13 M phenylmethylsulfonyl fluoride (PMSF), 0.03 M N-alpha-tosyl-L-lysyl-chloromethyl ketone (TLCK), and 0.03 M N-tosyl-L-phenylalanyl-chloromethyl ketone (TPCK). The bacterial suspensions were disrupted by Emulsiflex C3 high-pressure homogenizer (Avestin) operated at a range of 15,000-20,000 psi for 7 times at 4 °C. The lysates were centrifuged at 30,000 x g at 4 °C for 1 hr to separate insoluble and soluble proteins by using a Hitachi himac CP80WX ultracentrifuge (Hitachi Koki Co. Ltd., Tokyo, Japan). Then, 5 mL supernatant containing the soluble proteins were loaded onto a 1-ml MBPTrap HP column (GE Healthcare Bio-Sciences), which was pre-equilibrated with 20 mM Tris-HCl, pH 7.4, 200 mM NaCl, 1 mM EDTA, and 1 mM DTT, at a flow rate of 0.5 ml/min at 4 °C by ÄKTA Purifier. The column was eluted with 20 mM Tris-HCl, pH 7.4, 200 mM NaCl, 1 mM EDTA, 1 mM DTT, and 10 mM maltose at a flow rate of 1 ml/min at 4 °C by ÄKTA Purifier. The flow-through and elution fractions were analyzed by SDS-PAGE on a 12% gel. The elution fractions containing recombinant MBP-tagged HP-NAPR77-E116 were collected and concentrated to a concentration higher than 1 mg/ml.

*Recombinant HP-NAP-based enzyme linked immunosorbent assay (ELISA)*

Nunc MaxiSorp ninety-six-well enzyme linked immunosorbent assay (ELISA) plates (Nunc, Rochester, NY, USA) were coated with 0.3 μg of recombinant HP-NAP in 100 μl of bicarbonate buffer (pH 9.0) for each well at room temperature for 16 hr. Each well was washed with 300 μl of phosphate buffered saline (PBS), pH 7.4, containing 20 mM Na2HPO4, 1.47 mM KH2PO4, 137 mM NaCl, and 2.7 mM KCl, with the addition of 0.1% tween-20 (PBS-T) three times for 10 min each time. The wells were blocked with 250 μl PBS with 1% bovine serum albumin (BSA) for 2 hr and then washed with 300 μl of PBS-T buffer three times for 10 min each time. The anti-FLAG M2 antibody (Sigma-Aldrich, Cat# F-3165) and its corresponding mouse IgG antibody (Sigma-Aldrich, Cat# I5381) at a concentration of 600 ng/ml and the hybridoma culture supernatant containing mouse monoclonal antibody MAb 16F463 against HP-NAP at a dilution of 1:5000 in 100 μl of PBS-T buffer containing 1% BSA were added into each well. The plate was incubated at room temperature for 1 hr and then the wells were washed with 300 μl of PBS-T buffer three times for 10 min each time. The horseradish peroxidase-conjugated goat anti-mouse secondary antibody (Jackson ImmunoResearch) at a dilution of 1:10000 in 100 μl of PBS-T buffer containing 1% BSA was loaded into each well. The plate was incubated at room temperature for 1 hr and then the wells were washed with 300 μl of PBS-T buffer three times for 10 min each time. The color was developed using 3,3’,5,5’-tetramethylbenzidine (TMB) peroxidase substrate (Thermo Scientific). The reaction was terminated by the addition of 2 N H2SO4, and the absorbance at 450 nm was measured by Bio-rad iMark microplate absorbance reader (Hercules, CA).

*Western blot analysis*

Western blotting was performed essentially the same as previously described64. The membrane was probed with either anti-FLAG M2 antibody (Sigma-Aldrich) at a concentration of 1 μg/ml or the hybridoma culture supernatant containing mouse monoclonal antibody MAb 16F463 against HP-NAP at a dilution of 1:2000.

*Finding PP2A-like Peptides*

Knowing that the K374, G378, and Y381 of PP2A plays important roles in PP2A's favourable interactions with Sgo1 (see **Fig S4**), we set out to search for a peptide with a similar amino acid pattern and which may interact better than with Sgo1 and thereby out compete/displace PP2A and bind to Sgo1. Therefore, we searched the developed TP-DB for the following queries: **[K 3 G 2 Y]**, **[A 3 K 3 G], [G 2 Y 3 A], [A 3 K 3 G 2 Y], [K 3 G 2 Y 3 A], [A 3 K 3 G 2 Y 3 A], and [A 2,3 K 2,3 G 2,3 Y 2,3 A].**  The resulting peptides from each query are ranked based on their helical propensity score and contact number.

*Minimization and MD simulations of Martini Coarse-grained (CG) models in vacuum and assessing the stability of complexes*

To search for PP2A-like helical peptides to block the interaction between Sgo1 and PP2A, 69 sequences (<https://dyn.life.nthu.edu.tw/design/result?JobID=602bae01v>) that matched the pattern “K\*\*\*G\*\*Y” in PP2A helix (see **Fig S4**; the three anchoring residues of PP2A helix **K**374**T**375**I**376**H**377**G**378**L**379**I**380**Y**381 that interacts with Sgo1 are K374, G378 and Y381) were chosen for affinity assessment. Methodologically, it was equivalent to computationally modify (mutate) the five spacing residues T375, I376, H377, L379 and I380 in PP2A helix “KTIHGLIY” into corresponding residues in every of these 69 sequences. In other words, we only replaced the alpha-helical segment “K374TIHGLIY381” in PP2A [PDB:3FGA; chain B] by every of the 69 helical stretches that matched the pattern “K\*\*\*G\*\*Y” and the segment was shown to interact with the helix “V75KEAQDIILQLRKECYYL92” in Sgo1 [PDB:3FGA; chain D]. Subsequently, all 69 complexes comprising the Sgo1 helix and mutated PP2A helix were converted into coarse-grained (CG) models by the web service CHARMM-GUI65. Both terminals of the helical peptides in all complexes were set to be electrically neutral, and 69 CG models with PP2A mutants as well as the counterpart of wide-type were energy-minimized and then briefly equilibrated for 1 ps at 20 fs time step by MD simulations running on GROMACS package66 with MARTINI force field67, 68. The pressure and temperature were maintained at 1 bar and 310K, respectively. The cutoff distances for both Coulomb and Van der Waal were set to be 12Å.

MD simulations were conducted in vacuum with modest positional restraints (spring constant = 2.5 kcal/mol/Å2) exerted on the backbone beads for 1 ps at 20 fs time step, and the potential energies at the final step were recorded for the analysis to obtain the binding stability between the targeted helix in Sgo1 and mutated (or wild-type) PP2A segments.

Besides the top-ranked peptides by Martini forcefield from the aforementioned 69 complexes, we also searched potential helical binders from TP-DB using at least 4 anchoring residues in the search pattern (**Tables S5** to **S8**). We carried out all-atom molecular dynamics simulations for these peptides to assess their interactions with the targeted helix of Sgo1 using MM/PBSA69, 70 (see below). The top 11 identified helices (**Table S9**), searched from the queries [A 3 K 3 G 2 Y], [K 3 G 2 Y 3 A] and [A 2,3 K 2,3 G 2,3 Y 2,3 A], were first superimposed onto the template helix of the PP2A [PDB: 3FGA; chain B; residue 372 to 383] with their common residues K374, G378, and Y381 (as shown in Fig S3), before the simulations.

We carried out the all-atom molecular dynamics simulations by AMBER1671, 72 package, using ff14SB forcefield73 for protein and ionsjc\_tip3p for ions74 in explicit solvent. Each molecular system to be simulated was placed in a periodic box and solvated with TIP3P water. All the input files for MD simulations were prepared using tLeap71 from AmberTools16.

Energy minimization was performed in three stages - first, with weak harmonic positional restraints (spring constant = 0.5 kcal/mol/Å2) on all atoms except for the water/solvent atoms; second, with weak harmonic restraints on the CA atoms of amino acids; lastly, without restraints. Each of the first two stages of energy minimization is composed of 5,000 steps, with the first 2500 steps (for each energy minimization stage) carried out using steepest descent algorithm and the remaining 2,500 steps carried out using conjugate gradient algorithm.

Following the energy minimization, each simulation system was slowly heated to 310K while applying weak harmonic positional restraints on the CA atoms. Each system was then equilibrated at the temperature reached without any positional restraints prior to production MD simulations. Each production MD simulation ran for 500 ns at 2 fs time step. Non-bonded interactions were evaluated up to a cut-off distance of 10 Å where it was switched off with a cubic spline switch function. Particle Mesh Ewald method23 was used to calculate full electrostatic interactions energies. All temperature regulations were done using Langevin thermostat (with a collision frequency, γ, of 2 ps-1) and all pressure control was done with Berendsen barostat75.

Binding free energy change (binding ∆G) was calculated from the MD simulations trajectories using AMBER16’s implementation of Molecular Mechanics/Poisson-Boltzmann Surface Area (MM/PBSA)69, 70.

Supplementary References

55. Salic A, Waters JC, Mitchison TJ. Vertebrate shugoshin links sister centromere cohesion and kinetochore microtubule stability in mitosis. *Cell* **118**, 567-578 (2004).

56. Wang L-H, Yen C-J, Li T-N, Elowe S, Wang W-C, Wang LH-C. Sgo1 is a potential therapeutic target for hepatocellular carcinoma. *Oncotarget* **6**, 2023 (2015).

57. Liu H, Rankin S, Yu H. Phosphorylation-enabled binding of SGO1–PP2A to cohesin protects sororin and centromeric cohesion during mitosis. *Nature cell biology* **15**, 40-49 (2013).

58. Tang Z, Shu H, Qi W, Mahmood NA, Mumby MC, Yu H. PP2A is required for centromeric localization of Sgo1 and proper chromosome segregation. *Developmental cell* **10**, 575-585 (2006).

59. Marston AL. Shugoshins: tension-sensitive pericentromeric adaptors safeguarding chromosome segregation. *Molecular and cellular biology* **35**, 634-648 (2015).

60. Wang C-A, Liu Y-C, Du S-Y, Lin C-W, Fu H-W. Helicobacter pylori neutrophil-activating protein promotes myeloperoxidase release from human neutrophils. *Biochemical and biophysical research communications* **377**, 52-56 (2008).

61. Yang Y-C*, et al.* High yield purification of Helicobacter pylori neutrophil-activating protein overexpressed in Escherichia coli. *BMC biotechnology* **15**, 1-11 (2015).

62. Walker IH, Hsieh P-c, Riggs PD. Mutations in maltose-binding protein that alter affinity and solubility properties. *Applied microbiology and biotechnology* **88**, 187-197 (2010).

63. Iankov ID, Haralambieva IH, Galanis E. Immunogenicity of attenuated measles virus engineered to express Helicobacter pylori neutrophil-activating protein. *Vaccine* **29**, 1710-1720 (2011).

64. Hong Z-W, Yang Y-C, Pan T, Tzeng H-F, Fu H-W. Differential effects of DEAE negative mode chromatography and gel-filtration chromatography on the charge status of Helicobacter pylori neutrophil-activating protein. *PloS one* **12**, e0173632 (2017).

65. Lee J*, et al.* CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. *Journal of chemical theory and computation* **12**, 405-413 (2016).

66. Pronk S*, et al.* GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **29**, 845-854 (2013).

67. de Jong DH*, et al.* Improved parameters for the martini coarse-grained protein force field. *Journal of chemical theory and computation* **9**, 687-697 (2013).

68. Wassenaar TA, Ingólfsson HI, Böckmann RA, Tieleman DP, Marrink SJ. Computational lipidomics with insane: a versatile tool for generating custom membranes for molecular simulations. *Journal of chemical theory and computation* **11**, 2144-2155 (2015).

69. Kollman PA*, et al.* Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Accounts of chemical research* **33**, 889-897 (2000).

70. Wang J, Hou T, Xu X. Recent advances in free energy calculations with a combination of molecular mechanics and continuum models. *Current Computer-Aided Drug Design* **2**, 287-306 (2006).

71. Salomon‐Ferrer R, Case DA, Walker RC. An overview of the Amber biomolecular simulation package. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **3**, 198-210 (2013).

72. Salomon-Ferrer R, Götz AW, Poole D, Le Grand S, Walker RC. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh Ewald. *Journal of chemical theory and computation* **9**, 3878-3888 (2013).

73. Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *Journal of chemical theory and computation* **11**, 3696-3713 (2015).

74. Joung IS, Cheatham III TE. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *The journal of physical chemistry B* **112**, 9020-9041 (2008).

75. Berendsen HJ, Postma Jv, van Gunsteren WF, DiNola A, Haak JR. Molecular dynamics with coupling to an external bath. *The Journal of chemical physics* **81**, 3684-3690 (1984).