Organic & Biomolecular Chemistry

Correlations between secondary structure- and protein-protein interface-mimicry: The Interface Mimicry Hypothesis

Journal:	*Organic & Biomolecular Chemistry*
Manuscript ID	OB-ART-01-2019-000204.R1
Article Type:	Paper
Date Submitted by the Author:	26-Feb-2019
Complete List of Authors:	Taechalertpaisarn, Jaru; Texas A&M University, Chemistry Lyu, Rui-Liang; Texas A&M University System, Chemistry Arancillo, Maritess; TAMU, Chemistry Lin, Chen-Ming; TAMU, Chemistry Perez, Lisa; Texas A&M University, Chemistry Ioerger, Thomas ; Texas A&M University, Computer Science Burgess, K.; Texas A&M University Department of Biochemistry and Biophysics,

Correlations between secondary structure- and protein-protein interface-mimicry: The Interface Mimicry Hypothesis

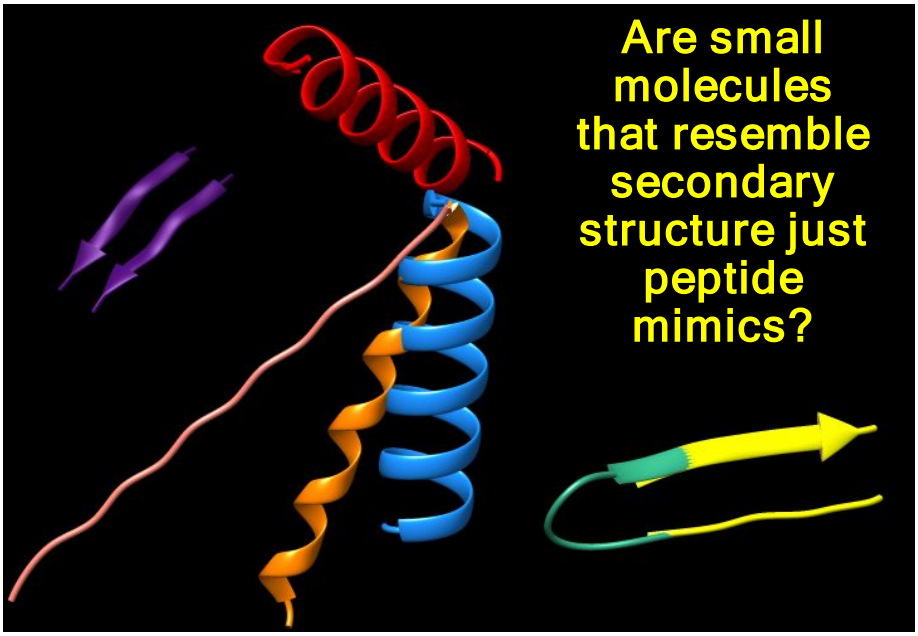
Jaru Taechalertpaisarn,[a] Rui-Liang Lyu,[a] Maritess Arancillo,[a] Chen-Ming Lin,[a] Lisa M. Perez,[b] Thomas R. Ioerger,[c] and Kevin Burgess[a]*

[a]*Department of Chemistry and* [b]*Laboratory For Molecular Simulation, Texas A & M University, Box 30012, College Station, TX 77842-3012, USA.* [c]*Department of Computer Science, Texas A & M University, College Station, TX 77843-3112.*

E-mail: burgess@tamu.edu

TITLE RUNNING HEAD: Secondary structure and protein-protein interface mimicry.

GRAPHICAL ABSTRACT



ABSTRACT

An active segment of the research community designing small molecules ("minimalist mimics" of peptide fragments) to interfere with protein-protein interactions have based their studies on an implicit hypothesis. Here we refer to this as the Secondary Structure Hypothesis, that might be defined as, "If a small molecule can orient amino acid side-chains in directions that resemble side-chains of the parent secondary structure at the interface, then that small molecule is a candidate to perturb the protein-protein interaction." Rigorous tests of this hypothesis require co-crystallization of minimalist mimics with protein receptors, and comparison of the bound conformations with the interface secondary structures they were designed to resemble. Unfortunately, to the best of our knowledge, there is no such analysis in the literature, and it is unlikely that enough examples will emerge in the near future to test the hypothesis. Research described here was designed to challenge this hypothesis from a different perspective. In a previous study, preferred conformations of a series of novel minimalist mimics were simulated then systematically overlaid on >240,000 crystallographically characterized protein-protein interfaces.[1] Select data from that overlay procedure revealed chemotypes that overlay side chains on various PPI interfaces with a relatively high frequency of occurrence. The first aim of this work was to determine if good secondary structure mimics overlay frequently on PPI interfaces. The second aim of this work was to determine if overlays of preferred conformers at interface regions involve secondary structures. Thus situations where these conformations overlaid extremely well on PPI interfaces were analyzed to determine if secondary structures featured the PPI regions where these molecules overlaid in the previous study. Combining conclusions from these two studies enabled us to formulate a hypothesis that is complementary to the Secondary Structure Hypothesis, but, unlike this, is supported by abundant data. We call this the Interface Mimicry Hypothesis.

INTRODUCTION

Many protein-protein interfaces feature secondary structures. The *Secondary Structure Hypothesis* is based on the supposition that if a small molecule can orient amino acid side-chains in directions that resemble side-chains of the parent secondary structure at the interface, then that small molecule is a candidate to perturb the protein-protein interaction (PPI). This hypothesis provides a logical approach to the design of small molecule probes and pharmaceutical leads involving PPIs, in an area where so few design criteria have been identified. It also appears to be valuable because there are numerous examples where it has been used to design minimalist mimics (small molecules presenting amino acid side-chains)[2, 3] that do, in fact, disrupt PPIs (specific cases,[4-9] reviews[10-15]). However, this is circumstantial evidence in support of the hypothesis, and there are few ways to definitively prove it; one is via X-ray analysis of minimalist mimics co-crystallized with their protein receptors. This would allow comparison of bound conformations with the secondary structure in protein ligand at the PPI interface. This strategy would reveal if the molecule binds the intended receptor region, *and* if the small molecule also mimics the ligand secondary structure that it was designed to. However, to the best of our knowledge, there is no analysis of this kind in the literature. Consequently, even though secondary structure mimicry is widely seen as a fast-track to molecules that disrupt specific PPIs, the value of secondary structure mimicry is assumed.

We thought it would be valuable to use a combination of molecular dynamics calculations and data mining to probe the Secondary Structure Hypothesis in a different way. To achieve this we used a combination of four computational approaches: EKO, EKOS, DSSP, and STRIDE.

EKO (*Exploring Key Orientations*)[3, 16] evaluates how preferred conformers of minimalist mimics align side-chains proteins at PPI interfaces. This strategy simulates thermodynamically preferred conformations of the small molecules and compares them with interface regions of PPIs; it can be performed on a huge scale by mining crystallographic data from PDB. If a small molecule *cannot* orient side chains in a similar way to the protein ligand, then it is *unlikely* to perturb the corresponding PPI by putting side chains in those orientations.

EKOS (*Exploring Key Orientations on Secondary structures*) is similar to EKO, except it compares preferred small molecule conformations *with ideal secondary structures*. If, for instance, there is not a preferred conformation of a small molecule that can present amino acids side chains in the same way as a targeted secondary structure, it cannot be an effective minimalist mimic of that structure. Both EKO and EKOS facilitate ways of screening out minimalist mimics that cannot be useful, and sometimes, when a good fit is found, they hint at ones that can. EKO does *not* rely on the Secondary Structure Hypothesis since it matches conformations of the molecules on PPI interfaces *without regard to secondary structure*. Conversely, EKOS *only* considers fit on secondary structures, but does not explore fits at PPI interfaces.

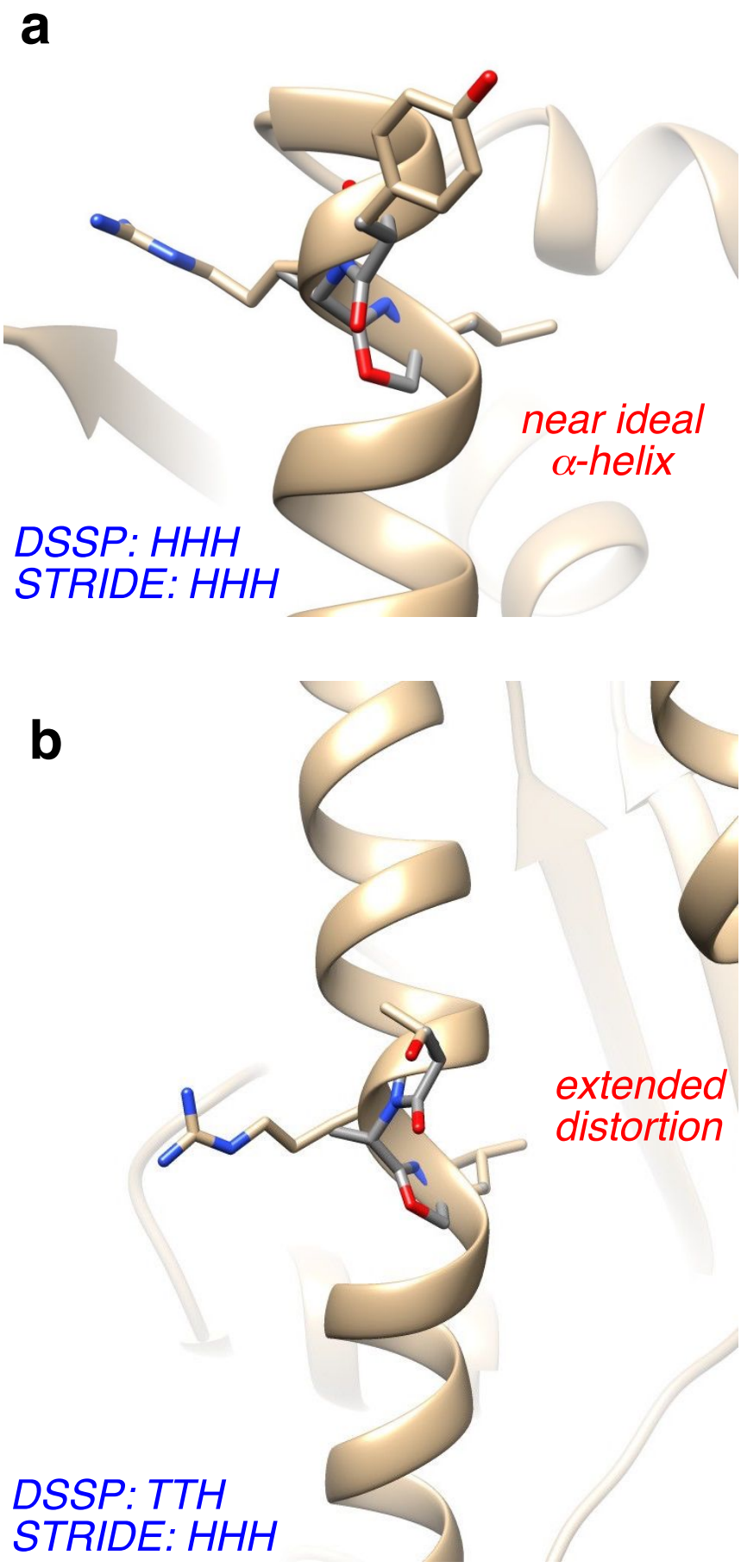
In the study described here, EKOS data was used to determine if preferred conformers of minimalist mimics overlay on all common secondary structures. These data are then compared with EKO processing of the same preferred conformers systematically overlaid on >240,000 PPI interfaces, as reported previously by us.[1] Here we required an algorithm to determine if the PPI overlay regions featured a secondary structure. This would facilitate determination of whether or not good mimics of ideal secondary structures actually tend to adopt those conformations at PPI interfaces; if not, perhaps the Secondary Structure Hypothesis should be modified or expanded for use as a predictive tool.

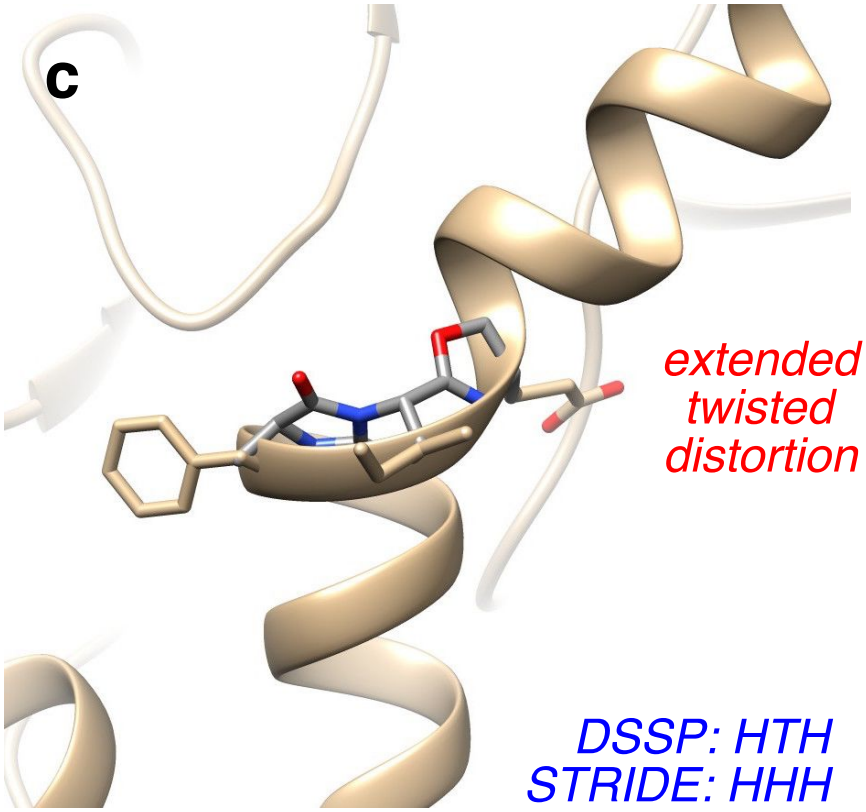
A program called *Dictionary of Secondary Structures of Proteins* (DSSP) was used to evaluate if the side-chains where preferred conformers matched PPI interfaces (from EKO) feature a secondary structure. DSSP identifies secondary structure motifs based on hydrogen-bonding patterns.[17, 18] To verify DSSP data, another program, *STRIDE*, was also used. Like DSSP, STRIDE evaluates protein residues in terms *H*-bonding patterns, but it also uses dihedral-angle parameters.[19, 20]

For EKOS, we considered the following nine ideal secondary structures: α-, π- and 3_{10}-helices; strand-turn-strands; regular and inverse γ-turns; β-strands; and, parallel and anti-parallel β-sheets. DSSP and STRIDE categorize parts of proteins into the following six types: α-, π- and

3_{10}-helices; strands, sheets, turns or bends, and those that do not have any recognizable secondary structures (which we refer to as “segments” later on for simplicity). Figure 1 shows typical data from combining EKO, DSSP and STRIDE. Here preferred conformations of a minimalist mimic overlaid on a protein ligand at an interface (as determined using EKO), and DSSP plus STRIDE were used to ascertain if this overlay region contains a secondary structure.

Throughout our analyses, a mimic that overlays only two of its three side-chain residues on an interface secondary structure was regarded as one that does not closely resemble it. Figure 1 features helical regions, but illustrates analyses is for any common secondary structure. Figure 1a shows an overlay on a near-ideal α-helical region, and both DSSP and STRIDE recognize that all three residues place side chains on that helix (HHH). In general, if the two programs are in agreement then this unambiguously identifies the overlay region as containing that secondary structure. If *neither* DSSP nor STRIDE recognize a secondary structure in the overlay region, then it was understood that segment unambiguously does not feature a secondary structure (*eg* Figure 1d). In Figures 1b and c, DSSP calls the overlay unstructured (a segment) while STRIDE assigns “helix”. After analyzing many overlays, we concluded that if DSSP and STRIDE are not in agreement, the assignment is truly borderline. Figures 2b and c were included to illustrate such cases where DSSP and STRIDE do *not* agree. Comparison of DSSP and STRIDE data throughout this text tends to indicate STRIDE tolerates more deviation from ideal than DSSP, *ie* DSSP tends to uphold higher standards before “calling” a secondary structure. However, the difference in DSSP and STRIDE outputs is not significant enough to affect the overall conclusions.


a
near ideal
α-helix
DSSP: HHH
STRIDE: HHH
b
extended
distortion
DSSP: TTH
STRIDE: HHH


c
extended
twisted
distortion
DSSP: HTH
STRIDE: HHH

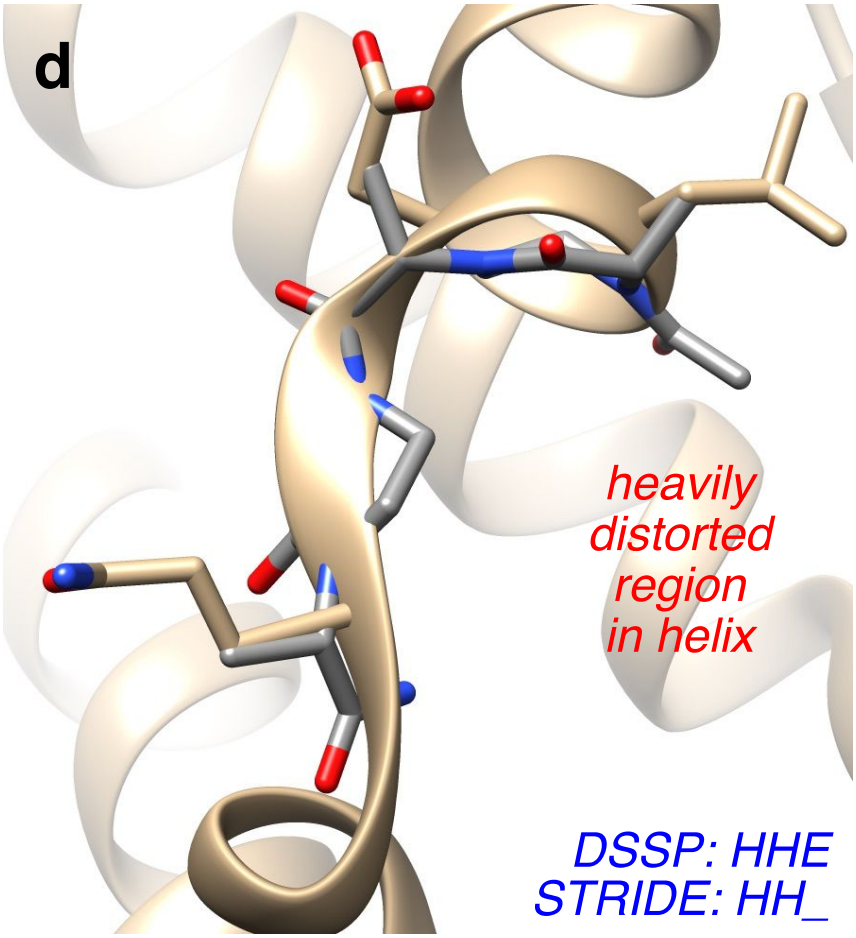
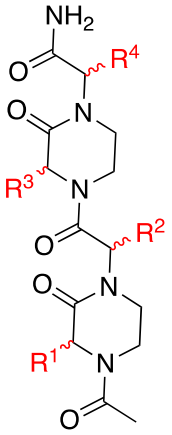
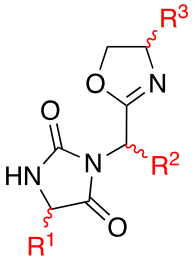

d
heavily
distorted
region
in helix
DSSP: HHE
STRIDE: HH_

Figure 1. Illustrative DSSP and STRIDE secondary structure assignments at protein interfaces where mimics overlay. **a** A mimic conformer overlays on a near-ideal α-helical fragment at a PPI interface; both DSSP and STRIDE recognize that region as helical (H). **b** A conformer is overlaid on an extended region between two helical segments that is hard to characterize; DSSP interprets the overlaid region as turn, turn, and helical (TTH), STRIDE calls them as uniformly helical (HHH), and visually we concluded that this overlay was ambiguous. **c** It is unclear whether the extended, twisted region shown is helical in this case, DSSP bins that as a segment, while STRIDE calls it as a helix. **d** A mimic overlaid on an extremely distorted region between two helical fragments, both DSSP and STRIDE bin this overlay as a segment, and we agree that the overlay is *not* on secondary structure.

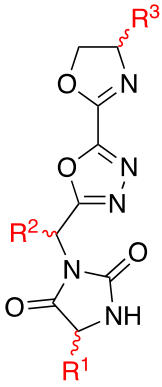
In a previous study we conceived eight new chemotypes that have not been reported previously, and compared their preferred conformations with >240,000 interfaces.[1] Four of these, **1** – **4**, overlaid much more frequently at interfaces than the others (not shown). Preferred conformers of **1** – **4** also overlaid on PPI interfaces more frequently Arora's oxopiperazine chemotype **A**[21] (included as a reference). In actuality, the fit of these preferred conformers followed the trend **1** >> **2** > **3** > **4** > **A**, where **1** was a far better interface mimic than the others. Thus the first specific aim of the work was to determine if good secondary structure mimics overlay frequently on PPI interfaces. The second specific aim was to analyze superior interface overlays from EKO[1] for the presence of secondary structures (as determined by DSSP and STRIDE).



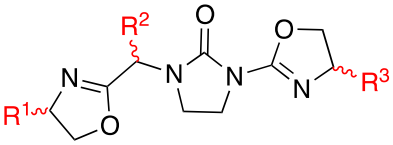
A



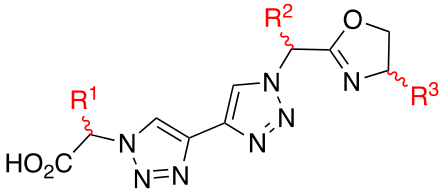
1



2



3



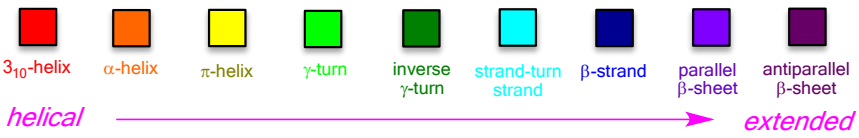
4

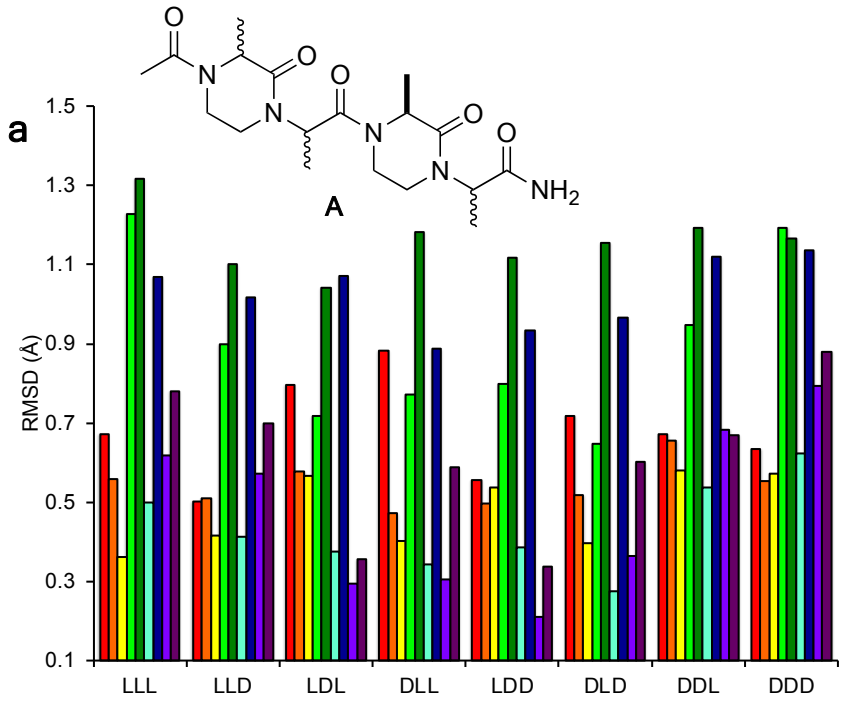
RESULTS AND DISCUSSION

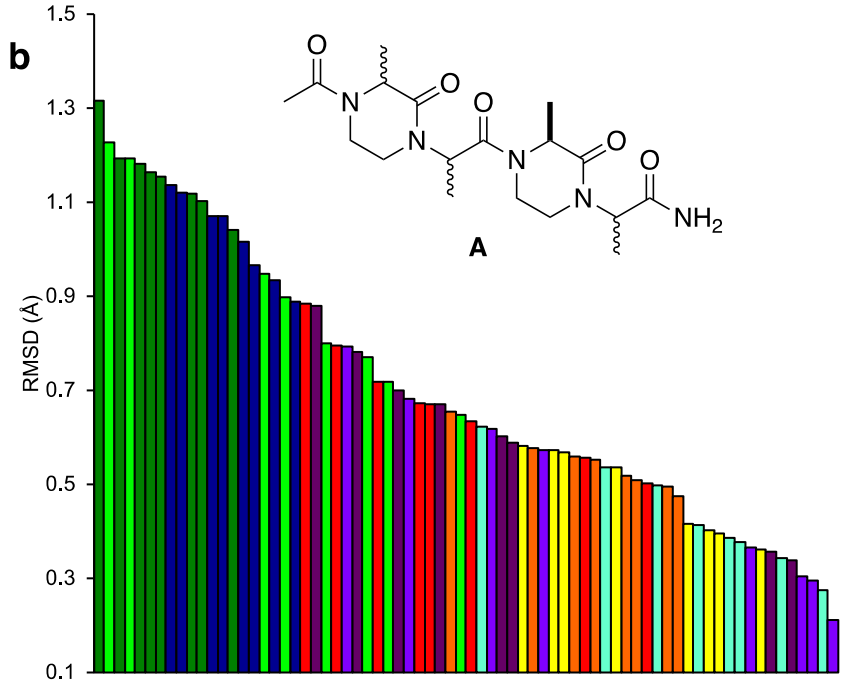
Peptidomimetic A

EKOS analyses of the trimethyl-substituted chemotype LLL-**Aaaa** ("**aaa**" denotes three methyl side chains analogous to AlaAlaAla; the **aaa** nomenclature is often omitted in this paper for simplicity) indicates it tends to overlay select common secondary structures better than the most effective minimalist mimics as of 2014.[3] Consequently, chemotype **A** is a useful benchmark for good interface mimic design.

Data from an EKOS analysis featuring *all* the isomers of **Aaaa** were obtained in the current study, whereas the original report[21] only featured the LLL-isomer. Figure 2a shows how each of the eight possible stereoisomers (grouped on the x-axis) overlay on the ideal secondary structures, and 2b arranges the best matching conformers in descending RMSD of the overlays irrespective of stereochemistry. The best overlay identified was for LDD-**A** on a parallel β-sheet (RMSD 0.21 Å). Figure 2c illustrates that best fit; the orientations of the side-chains in the ideal parallel β-sheet and the simulated conformer are indeed very close.







best matches of **A** stereoisomers

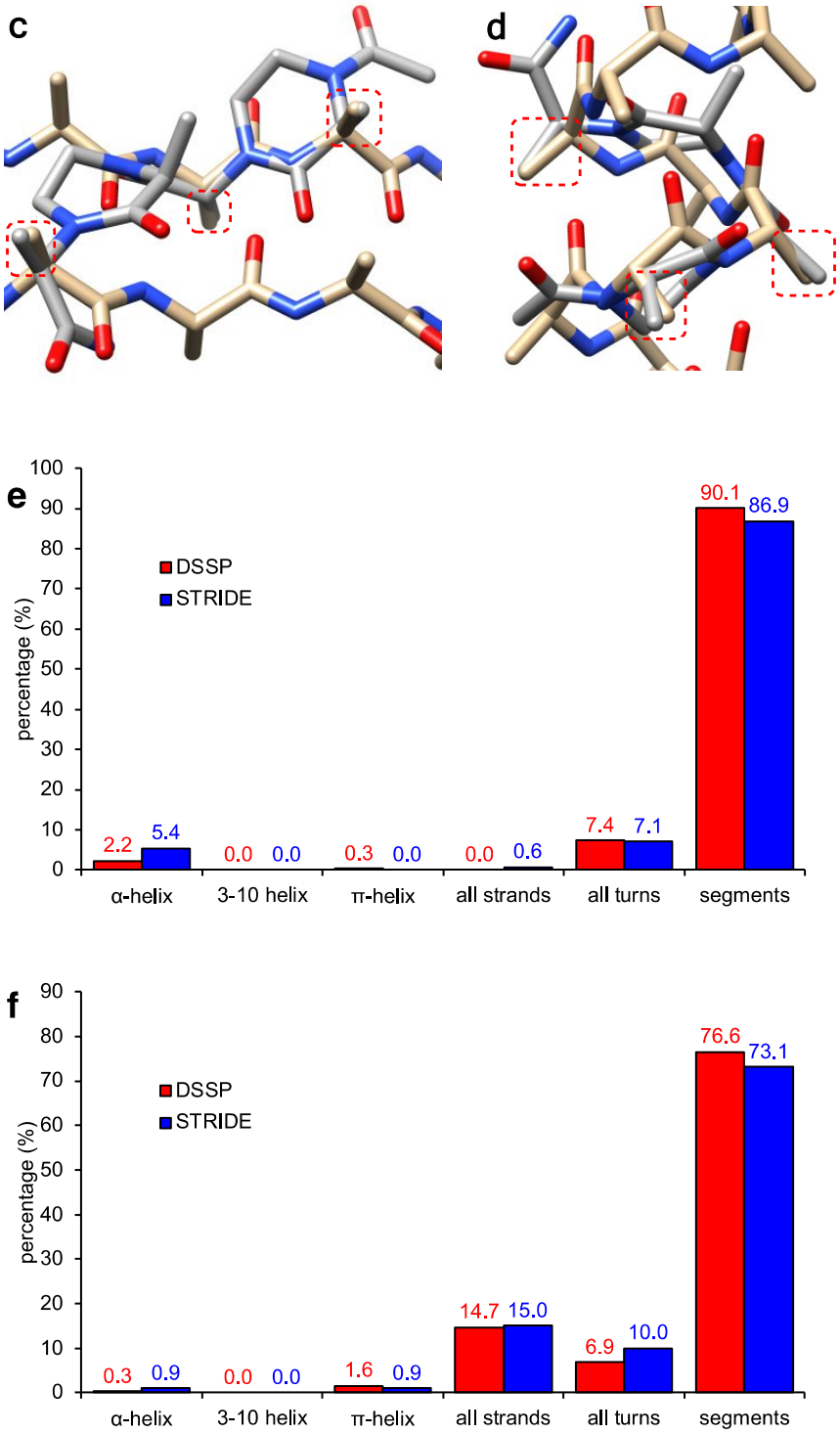


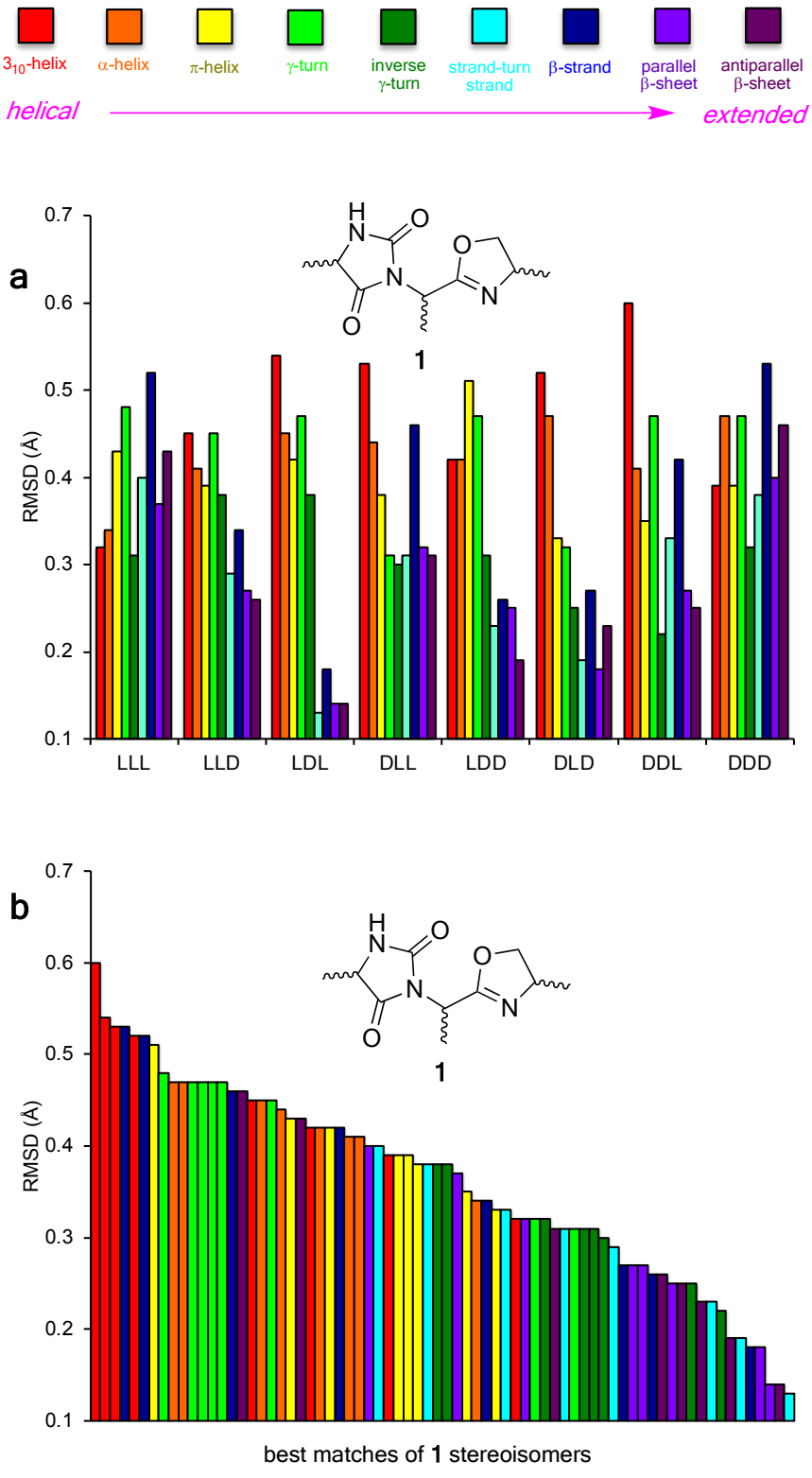
Figure 2. RMSD (Å) of the overlays of mimics **A** on each of the ideal secondary structures, organized by stereochemistry (**a**) or by decreasing RMSD (**b**). Overlay of preferred conformers of LDD-**A** (silver) on a parallel β-sheet (gold), RMSD 0.21 Å (**c**); and, of LLL-**A** on a π-helix (also gold), RMSD 0.36 Å (**d**). Statistical distribution of secondary structures at PPI interfaces derived by DSSP and STRIDE calculations; (**e**) the best 312 overlays of LLL-**A** (all RMSDs < 0.25 Å); and, (**f**) 320 overlays of LDD-**A** (RMSD < 0.25 Å). Note that calculations do not differentiate strand-turn-strand, parallel- and antiparallel-sheets.

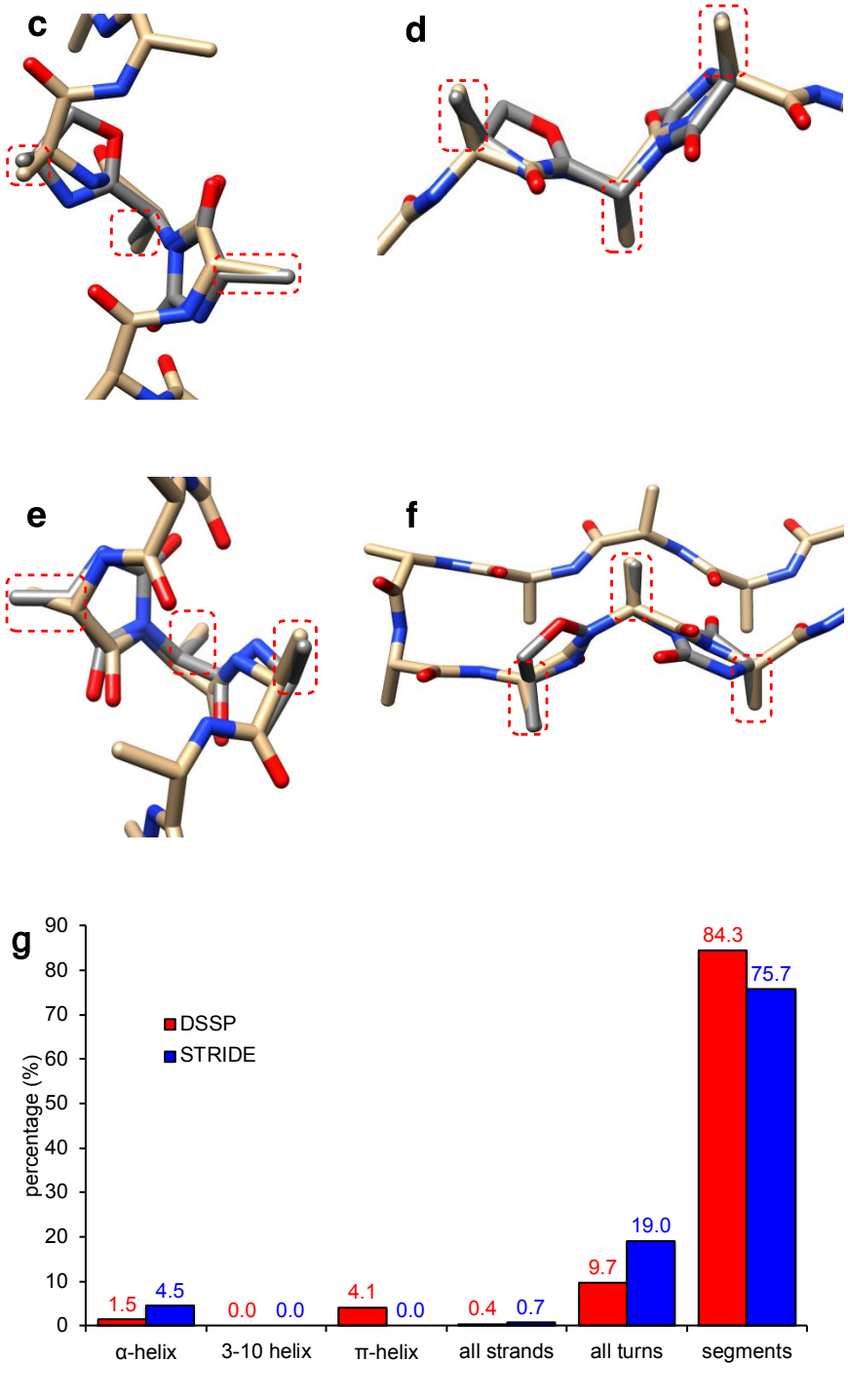
Figure 2a reveals LLL-**A** is a good mimic for helices, and LDD-**A** is better at mimicking extended structures. Consequently, it seemed likely that LLL-**A** would overlay more frequently on helices at PPI interfaces in the PDB, and LDD-**A** would overlay well more frequently on strands and sheets. To check if this is true, we selected the best overlays for each stereoisomer (RMSD <0.25 Å based the three side-chains) from our previous EKO analysis on >240,000 PPI interfaces.[1] This approach generated 312 and 320 PPI interface matches for LLL- and LDD-isomer, respectively. Each match was then analyzed using the DSSP and STRIDE programs. To our surprise, only a small portion of these matches was on regions with clear secondary structures at all (Figure 2e and f). DSSP and STRIDE analyses indicate *most (>73 %) of the matches were on segments* (Figure 2f). Consistent with the Secondary Structure Hypothesis, LLL-**A** does in fact overlay more frequently on helices than LDD-**A** (2.2 and 0.3% of the overlays, as determined by DSSP), while LDD-**A** more frequently matches well on sheets and strands (14.7 and 0 %), but this only accounts for small fractions of the best overlays in each case.

Interface Mimics 1 – 4

Using exactly the same strategy as above, preferred conformers of mimic **1** were systematically overlaid on ideal secondary structures using EKOS. Figure 3b replots the data in Figure 3a, but from highest to lowest RMSD, irrespective of secondary structure. This presentation reveals **1** is a superior mimic compared with **A**. Chemotype **A** (Figure 2b; note the expansion of the y-axis showing RMSD, is different) only overlays well on antiparallel, parallel β-sheets, and strand-turn-strand secondary structures with RMSD <0.35 Å. Chemotype **1** is therefore an outstanding of ideal secondary structures. Our previous work showed structure **1** gave significantly more matches on PPI interfaces than **A** (over 180,000 matches for **1** compared to ~3,000 for **A**). These two sets of data combined show that good matches on secondary structures implies good overlays on PPI interfaces, just as observed for **A**. Most of the preferred conformers of LLL- and LDL-**1** that matched on PPI interfaces (75 and 56 %, respectively) did so on interface regions that did *not* on ideal secondary structures (Figure 3g-h).

Chemotypes **2**, **3**, and **4** were analyzed using exactly the same strategy as outlined above for **1** and **A**. Data for these experiments are shown in the Supporting Information. The high-level trend from this data is clear: **2**, **3**, and **4** (in that order) are fine secondary structure mimics, are all better than **A**, and not all are many times inferior to **1**. This is exactly the same trend observed in our previous work on overlaying their preferred conformations on >240,000 PPI interfaces. Analysis of DSSP/STRIDE data also revealed similar trend found with **1** and **A**, where the majority of matches by **2**, **3**, and **4** were on “segments”, despite they were fine secondary structure mimics.


3₁₀-helix
α-helix
π-helix
γ-turn
inverse γ-turn
strand-turn strand
β-strand
parallel β-sheet
antiparallel β-sheet
helical
extended
a
RMSD (Å)
0.7
0.6
0.5
0.4
0.3
0.2
0.1
LLL
LLD
LDL
DLL
LDD
DLD
DDL
DDD
1
b
RMSD (Å)
0.7
0.6
0.5
0.4
0.3
0.2
0.1
1
best matches of 1 stereoisomers


c
d
e
f
g
DSSP
STRIDE
percentage (%)
0
10
20
30
40
50
60
70
80
90
1.5
4.5
0.0
0.0
4.1
0.0
0.4
0.7
9.7
19.0
84.3
75.7
α-helix
3-10 helix
π-helix
all strands
all turns
segments

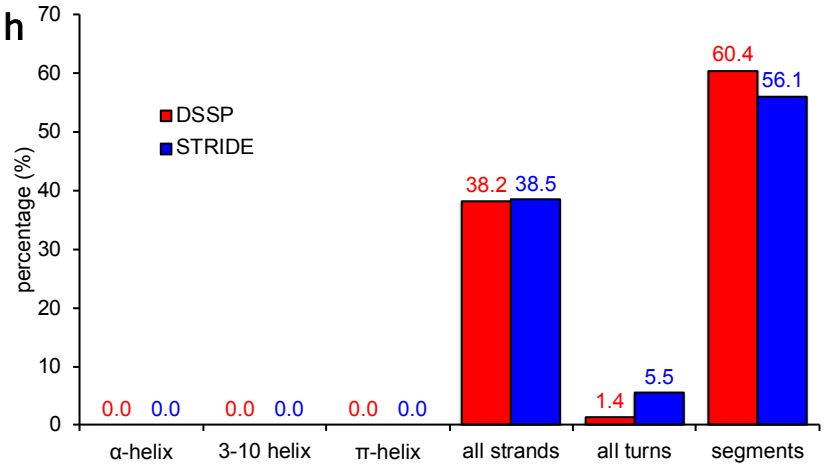


Figure 3. **a** Overlay data for the best matching accessible conformer of mimics **1** on each of the ideal secondary structures; **b** data in **a** replotted in descending RMSD (left to right) irrespective of stereochemistry. **c** – **f** Optimal overlays for low energy conformers of chemotype **1** (silver) on 3_{10}-helix (**c**), β-strand (**d**), α-helix (**e**), and strand-turn-strand (**f**; all in gold). The fit is perceptibly superior for **d** and **f**, but it is still close in **c** and **e**. Statistical distribution of secondary structures at PPI interfaces derived by DSSP and STRIDE calculations; (**g**) the best 268 overlays of LLL-**1** (all RMSDs < 0.15 Å); and, (**h**) 1008 overlays of LDL-**1** (< 0.10 Å RMSD).

CONCLUSIONS

The first aim of this work was to determine if secondary structure mimicry is a good predictor of interface mimicry, *ie* if good secondary structure mimics overlay frequently on PPI interfaces. Data in Figures 2, 3 and S1–S3, reveal that the relative potential of chemotypes **A**, **1**, **2**, **3** and **4** for secondary structure mimicry corresponds *exactly* to their tendency for interface mimicry as determined in our previous study *(ie **1** >> **2** > **3** > **4** > **A**)*.[1] Thus, besides being found frequently at PPI interfaces, chemotypes **1**, **2**, **3** and **4** are superior minimalist mimics of secondary structures, and **1** is truly exceptional again. Thus, to address the first aim of this work, good secondary structure mimics do, in fact, overlay frequently on PPI interfaces.

The second aim of this work was to determine if overlays of preferred conformers at interface regions involve secondary structures. In the event, *overlays on unstructured segments predominated* for every stereoisomer of each chemotype examined, without exception; in fact, there were only a few instances for which a bias towards any secondary structure represents over 30 % of the top hits. Thus, the conclusion for this aim is that particular preferred conformers of the minimalist mimics that overlay well on PPI interfaces do not tend to do so on secondary structure interface motifs; instead they overlaid far more frequently on interface regions that do not comprise a secondary structure.

Combination of the two conclusions described above indicates an interesting area for future research. For several decades, minimalist mimics have been evaluated for their potential to disrupt PPIs based on their tendency to be α-helical, β-turn, or sheet mimics, *etc*. (for reviews[11, 14, 22-32]). The implicit assumption is that if the corresponding secondary structure is found at a PPI interface, researchers would prioritize synthesis and testing of the corresponding minimalist mimic. However, as we already noted, there is little data on minimalist mimics co-crystallized with the protein receptors they were designed to bind to, and none where the bound conformations have been compared with the interface secondary

structures the compound was designed to resemble. A researcher may design a mimic of a secondary structure at a PPI interface, and observe experimentally that it does bind that protein receptor, but would still not know the bound conformation. However, it could be that the compound has affinity because it is a better mimic of protein segments in general, and may not necessarily adopt a bound conformation that resembles the targeted secondary structure.

To the best of our knowledge the Secondary Structure Hypothesis has been described but no one has attempted to rigorously define it. In the introduction of this paper we defined it in the following way:

Secondary Structure Hypothesis: if a small molecule can orient amino acids side-chains in directions that resemble those of the parent secondary structure at the interface, then that small molecule is a candidate to perturb the protein-protein interaction.

This paper does not confirm or refute this hypothesis, but it does lead us to a complementary one that is strongly supported by the huge amount of data processed in this study:

Interface Mimicry Hypothesis: small molecules that can orient amino acid side chains in directions that resemble secondary structures in general tend to be good *interface* mimics because they generally represent shapes of protein regions well.

Up until now, users of The Secondary Structure Hypothesis would have been constrained by the idea that it was only useful for PPIs that feature a secondary structure at the interface. The Interface Mimicry Hypothesis teaches minimalist mimics having preferred conformations that resemble secondary structures well, also tend to be good interface mimics even at interfaces that do *not* feature a secondary structure. Thus the Interface Mimicry Hypothesis predicts minimalist mimics that resemble secondary structures well also frequently adopt conformations that overlay on interface regions *with no secondary structure*. Said differently, good secondary structure mimics are most valuable simply because they are good peptide mimics in general.

The conclusion formulated above is the most important one to emerge from this work, but the data shows many other interesting trends that were not discussed above, because to do so would detract from reaching that conclusion. Some highlights from that data are outlined here.

Chemotype **1** is an especially good mimic of extended conformations {Figure 3b where the following color scheme is used: strand-turn-strand (light blue), β-strand (navy blue), parallel and antiparallel β-sheets (light and dark violet)}, *cf* blue and violet bars are concentrated at the low RMSD end of the chart. Overlays of preferred conformers of **1** on more twisted helical structures (red, orange, yellow bars) occur at higher RMSDs. However, stereoisomers of chemotype **1** can be found to overlay on *any* of the ideal secondary structures with RMSD < 0.35 Å. The LLL-isomer of **1** proved to be a better α- and 3_{10}-helical mimic than any of the other chemotypes **2** – **4** and **A**, but it also tended to overlay even better on other secondary structures. Chemotype LDL-**1** is interesting insofar as it *does* match on sheet-type structures with a 38 % frequency, consistent with EKOS data which showed **1** is a superior strand/sheet mimic. Several DSSP and STRIDE analyses were performed for this study, but the data in Figure 3h is notable because it shows the highest bias among all the chemotypes towards any secondary structure relative to “segments”.

Neither **2** nor **3** showed an significant bias towards overlays on helical structures, even though their shapes are chiral and non-planar. However, chemotype **4**, which contains two planar and aromatic heterocycles, showed most bias towards helicity. Like most minimalist mimics,[33, 34] **4** populates conformers that resemble several secondary structures and some of these are not helical but extended.

Overall, structure **A** tends to overlay better on extended structures than the helical ones.[3] For any helical structure, the best overlay was for LLL-**A** on the *i, i* + 1, *i* + 3 side-chains of a π-helix (0.36 Å RMSD; Figure 2d). The dotted red boxes in these graphics highlight how the chemotype side-chains align with those on the secondary structures. This tendency of

different stereoisomers to favor different secondary structures, in fact, applies to all other chemotypes in the rest of the study as well.

It is easy to assume helical minimalist mimics are easy to design because so many papers claim to do this. On the contrary, our findings indicate it is difficult to design minimalist mimics that overlay on helical secondary structures in preference to all others: true helical minimalist mimics are harder to conceive than similar sheet-mimics. Overlays on sheets can occur along one strand, or on two residues in one strand and another in the second strand. A mimic that spans across the sheet may do so perpendicular to the two strands, or diagonally. For strand-turn-strand, an overlaying mimic might interact with one part on the turn-region, hence there is broad latitude in sheet mimicry. Conversely, to mimic helical structures a compound must prefer conformations that are twisted with the targeted screw sense; this is simply harder to arrange.

Analyses of the type featured in this work are as reliable as the computational methods involved. In our opinion, the main limitation of QMD occurs in cases where not enough conformational space was sampled (leading to the possibility that some matches might be missed). However, over a large number of simulations, and involving closely related stereomers, the overall conclusions relating to the *Interface Mimicry Hypothesis* are not likely to change due to missed hits. The more detailed conclusions outlined above may have to be adjusted, but we do not anticipate a large variation at this stage.

Ultimately, the *Interface Mimicry Hypothesis* outlined here may be tested. This will probably occur when hits from libraries of secondary structure mimics are co-crystallized with their targets.

ASSOCIATED CONTENT

Electronic Supplementary Information (ESI) available: EKO, EKOS, and DSSP/STRIDE data for compounds **2** – **4**.

AUTHOR INFORMATION

Corresponding Author

burgess@tamu.edu

Notes

The authors declare no competing financial interests.

ACKNOWLEDGEMENTS

Financial support for this project was provided by the National Science Foundation (CHE-1608009), National Institutes of Health (HL126346), The Robert A. Welch Foundation (A-1121), DoD BCRP Breakthrough Award (BC141561), and CPRIT (RP150559 and RP170144). We thank Dr. Lisa M. Perez for useful discussions. NMR instrumentation at Texas A&M University was supported by a grant from the National Science Foundation (DBI-9970232) and the Texas A&M University System.

REFERENCES

1. J. Taechalertpaisarn, R.-L. Lyu, M. Arancillo, C.-M. Lin, Z. Jiang, L. M. Perez, T. R. Ioerger and K. Burgess, *Org. Biomol. Chem.*, 2019, **17**, 908-915
2. H. Yin and A. D. Hamilton, *Protein Secondary Structure Mimetics as Modulators of Protein-Protein and Protein-Ligand Interactions, in Chemical Biology*, Wiley-VCH-VCH Verlag, 2008.
3. D. Xin, E. Ko, L. M. Perez, T. R. Ioerger and K. Burgess, *Org. Biomol. Chem.*, 2013, **11**, 7789-7801.
4. H. Yin, G.-i. Lee, H. S. Park, G. A. Payne, J. M. Rodriguez, S. M. Sebti and A. D. Hamilton, *Angew. Chem. Int. Ed.*, 2005, **44**, 2704-2707.
5. J. C. Fuller, R. M. Jackson, T. A. Edwards, A. J. Wilson and M. R. Shirts, *PLoS One*, 2012, **7**, e43253.
6. J. L. Yap, X. Cao, K. Vanommeslaeghe, K.-Y. Jung, C. Peddaboina, P. T. Wilder, A. Nan, A. D. MacKerell, W. R. Smythe and S. Fletcher, *Org. Biomol. Chem.*, 2012, **10**, 2928-2933.
7. P. Ravindranathan, T.-K. Lee, L. Yang, M. M. Centenera, L. Butler, W. D. Tilley, J.-T. Hsieh, J.-M. Ahn and G. V. Raj, *Nat. Commun.*, 2013, **4**, 1-11.
8. B. B. Lao, K. Drew, D. A. Guarracino, T. F. Brewer, D. W. Heindel, R. Bonneau and P. S. Arora, *J. Am. Chem. Soc.*, 2014, **136**, 7877-7888.
9. B. B. Lao, I. Grishagin, H. Mesallati, T. F. Brewer, B. Z. Olenyuk and P. S. Arora, *Proc. Natl. Acad. Sci.* , 2014, **111**, 7531-7536.
10. P. S. Farmer and E. J. Ariens, *Trends Pharmacol. Sci.*, 1982, **3**, 362-365.
11. M. W. Peczuh and A. D. Hamilton, *Chem. Rev.*, 2000, **100**, 2479-2494.
12. M. Eguchi and M. Kahn, *Mini-Rev. Med. Chem.*, 2002, **2**, 447-462.
13. A. B. Smith, III, A. K. Charnley and R. Hirschmann, *Acc. Chem. Res.*, 2011, **44**, 180-193.
14. M. K. P. Jayatunga, S. Thompson and A. D. Hamilton, *Bioorg. Med. Chem. Lett.*, 2014, **24**, 717-724.
15. A. E. Modell, S. L. Blosser and P. S. Arora, *Trends Pharmacol. Sci.*, 2016, **37**, 702-713.
16. D. Xin, L. M. Perez, T. R. Ioerger and K. Burgess, *Angew. Chem. Int. Ed.* , 2014, **53**, 3594-3598.
17. G. Touw Wouter, C. Baakman, J. Black, E. Krieger, G. Vriend, A. H. te Beek Tim and P. Joosten Robbie, *Nucleic Acids Res*, 2015, **43**, D364-D368.
18. W. Kabsch and C. Sander, *Biopolymers*, 1983, **22**, 2577-2637.
19. D. Frishman and P. Argos, *Proteins Struct., Funct., Genet.*, 1995, **23**, 566-579.
20. M. Heinig and D. Frishman, *Nucleic Acids Res.*, 2004, **32**, W500-W502.
21. P. Tosovska and P. S. Arora, *Org. Lett.*, 2010, **12**, 1588-1591.
22. Y. Che, B. R. Brooks and G. R. Marshall, *J. Comput. Aided Mol. Des.*, 2006, **20**, 109-130.
23. M. J. Perez de Vega, M. Martin-Martinez and R. Genzalez-Muniz, *Curr. Top. Med. Chem.*, 2007, **7**, 33-62.
24. I. Saraogi and A. D. Hamilton, *Chem. Soc. Rev.*, 2009, **38**, 1726-1743.
25. H. Yin, G.-I. Lee and A. D. Hamilton, *Drug Discovery Res.*, 2007, 281-299.
26. H. Yin and A. D. Hamilton, in *Chemical Biology: From Small Molecules to System Biology and Drug Design*, eds. S. L. Schreiber, T. M. Kapoor and G. Wess, Wiley-VCH, 2007, vol. 1, pp. 250-269.
27. J. M. Davis, L. K. Tsou and A. D. Hamilton, *Chem. Soc. Rev.*, 2007, **36**, 326-334.
28. J. Becerril, J. M. Rodriguez, I. Saraogi and A. D. Hamilton, *Foldamers*, 2007, 195-228.
29. S. Fletcher and A. D. Hamilton, *J. Royal. Soc., Interface*, 2006, **3**, 215-233.
30. H. Yin and A. D. Hamilton, *Angew. Chem., Int. Ed.*, 2005, **44**, 4130-4163.
31. N. T. Ross, W. P. Katt and A. D. Hamilton, *Philos. Trans. R. Soc., A*, 2010, **368**, 989-1008.
32. V. Azzarito, K. Long, N. S. Murphy and A. J. Wilson, *Nat. Chem.*, 2013, **5**, 161-173.
33. E. Ko, J. Liu and K. Burgess, *Chem. Soc. Rev.*, 2011, **40**, 4411-4421.
34. E. Ko, J. Liu, L. M. Perez, G. Lu, A. Schaefer and K. Burgess, *J. Am. Chem. Soc.*, 2011, **133**, 462-477.

graphical abstract shown above abstract

Correlations between secondary structure- and protein-protein interface-mimicry

Jaru Taechalertpaisarn,[a] Rui-Liang Lyu,[a] Maritess Arancillo,[a] Chen-Ming Lin,[a] Lisa M. Perez,[b] Thomas R. Ioerger,[c] and Kevin Burgess[a]*

[a]Department of Chemistry and [b]Laboratory For Molecular Simulation, Texas A & M University, Box 30012, College Station, TX 77842-3012, USA. [c]Department of Computer Science, Texas A & M University, College Station, TX 77843-3112.

E-mail: burgess@tamu.edu

TITLE RUNNING HEAD: Secondary structure and protein-protein interface mimicry.

GRAPHICAL ABSTRACT