Electronic Supplementary Material (ESI) for RSC Advances.

Supporting Information for

"Termini Capping of Metal-Poly-His Peptide Complexes Induce Formation of α-Helix"

Eyal Simonovsky[a,b] Henryk Kozlowski[c,*] and Yifat Miller[a,b,*]

[a]Department of Chemistry, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel.[b]Ilse Katz Institute for Nanoscale Science and Technology, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel. [c] Faculty of Chemistry, University of Wroclaw, 50-383 Wroclaw, Poland

* Author to whom correspondence should be addressed:

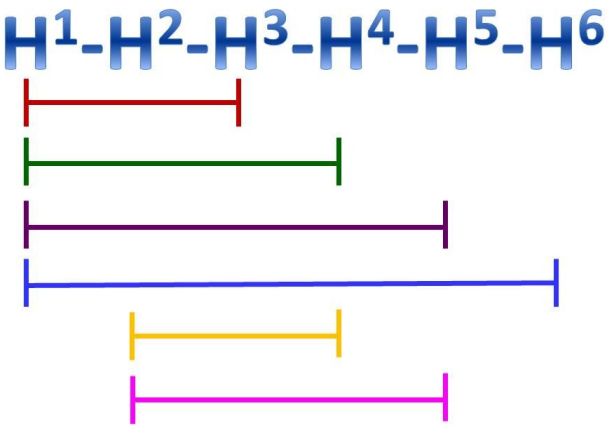
Yifat Miller: ymiller@bgu.ac.il

Henryk Kozlowski: henryk.kozlowski@chem.uni.wroc.pl

Materials and Methods

Constructed Cu^{2+}-His_6 structural models

The Cu^{2+} can bind to various binding sites in His_6. Previously, we have shown by applying experimental techniques that the Cu^{2+} binds two non-adjacent imidazole groups in His_6.[1] Considering the symmetry of His_6, i.e. the peptide consists of only His residues there is a total of six possible binding sites, as seen from scheme 1:



Scheme 1: The six possible binding sites that Cu^{2+} can bind to His_6.

We constructed the initial models using the Accelerys Discovery Studio software (http://accelrys.com/products/discovery-studio/). We interacted each His residue with Cu^{2+} for each model and accordingly we 'moved' the His residues closer to the Cu^{2+}. Applying parameterizations for the Cu^{2+}-peptide complex and constrains in the CHARMM force-field, we minimized the complex in aim to construct the initial predicted structure for each model. All His residues were modelled according to the physiological pH in all constructed peptide. The capping and uncapping of the termini had been performed using the CHARMM force-field.

Molecular dynamics (MD) simulations protocol

The 14 models were first minimized as previously we have performed for amyloids and other peptides.[2-8]

All-atom explicit MD simulations of the solvated models were performed in the NPT ensemble using NAMD[9] with the CHARMM27 force-field.[10,11] The models were energy minimized and explicitly solvated in a TIP3P water box[12,13] with a minimum distance of 15 Å from each edge of the box. Each water molecule within 2.5 Å of the models was removed. Counter ions were added at random locations to neutralize the models' charge. The Langevin piston method[14,15] with a decay period of 100 fs and a damping time of 50 fs was used to maintain a constant pressure of 1 atm. A temperature of 310 K was controlled by a Langevin thermostat with a damping coefficient of 10 ps.[9] The short-range van der Waals interactions were calculated using the switching function, with a twin range cut-off of 10.0 and 12.0 Å. Long-range electrostatic interactions were calculated using the particle mesh Ewald method with a cutoff of 12.0 Å.[16,17] The equations of motion were integrated using the leapfrog integrator with a step of 1 fs. The solvated systems were energy minimized for 2000 conjugated gradient steps. The counter ions and water molecules were allowed to move. The hydrogen atoms were constrained to the equilibrium bond using the SHAKE algorithm.[18] The minimized solvated systems were energy minimized for 5000 additional conjugate gradient steps and 20,000 heating steps at 250 K, with all atoms being allowed to move. Then, the system was heated from 250 K to 310 K for 300 ps and equilibrated at 310 K for 300 ps. All simulations were run for 100 ns at 310 K. Parameterizations for the Cu^{2+}-peptide complexes had been

performed for all MD simulations. The force constant values for Cu^{2+}-N atom are in the range of 10-50 kcal/mol/Å^2.

Generalized Born Method with Molecular Volume (GBMV)

The relative conformational energies can be compared only for models that have the same sequence and number of peptides, thus the relative conformational energies had been computed separately for the six capped and six uncapped models. The singly-capped (C- or N- terminal) peptide models C15 and D15 have different number of atoms, therefore cannot be compared with the other models (models A and B) or between each other.

To obtain the relative conformational energies of the twelve models of A and B, the models' trajectories of the last 5 ns were first extracted from the explicit MD simulations excluding the water molecules - a total of 500 conformations for each model. The solvation energies of all conformations were calculated using the GMBV.[19,20] In the GBMV calculations, the dielectric constant of water was set to 80. The hydrophobic solvent-accessible surface area (SASA) term factor was set to 0.00592 kcal/(mol Å). Each conformation was minimized using 1000 cycles, and the conformational energy was evaluated by grid-based GBMV.

A total of 3000 conformations (500 for each model) were used to construct the energy landscapes of the six capped and the six uncapped models and to evaluate the conformer probabilities by using Monte Carlo (MC) simulations. In the first step, one conformation of conformer i and one conformation of conformer j were randomly selected. Then, the Boltzmann factor was computed as $e^{-(Ej-Ei)/kT}$, where E_i and E_j are the conformational

energies evaluated using the GBMV calculations for conformations i and j, respectively, k is the Boltzmann constant and T is the absolute temperature (298 K used here). If the value of the Boltzmann factor was larger than the random number, then the move from conformation i to conformation j was allowed. After 1 million steps, the conformations 'visited' for each conformer were counted. Finally, the relative probability of model n were evaluate as $P_n = N_n/N_{total}$, where P_n is the population of model n, N_n is the total number of conformations visited for model n, and N_{total} is the total steps. The advantages of using MC simulations to estimate conformer probability lie in their good numerical stability and the control that they allow of transition probabilities among several conformers.

A total of 3000 conformations of the six capped peptide models (500 conformations for each model) and a total of 3000 conformations of the six capped peptide models (500 conformations for each model) were used to construct the energy landscape of the six capped peptide models and six uncapped peptide models (Table S1). The group of these twelve models is likely to present may be only a very small percentage of the ensemble. Nevertheless, the carefully selected models cover the most likely structures.

Assigning secondary structure to amino acids by the DSSP algorithm

The DSSP algorithm is the standard method for assigning secondary structure to the amino acids of a protein or a peptide, given the atomic-resolution coordinates of the protein or the peptide. It does this by reading the position of the atoms in a protein (the ATOM records in a PDB file) followed by calculation of the hydrogen bond energy between all atoms. The best two hydrogen bonds for each atom are then used to

determine the most likely class of secondary structure for each residue in the protein or the peptide. We applied the DSSP algorithm that is embedded in the CHARMM software.[21] The DSSP algorithm provides information on specific domains that illustrate α-helix and β-sheet .It does not distinguish between various types of α-helix, such as 3_{10} helix, α-helix, etc.

Analysis of the Cu^{2+}-Nε (His) atom and Cu^{2+}-O (C-terminal carboxyl group) atom distance distribution for each His residue using the JMP program

The vertical line within each box represents the median sample value. The ends of the box represent the 3rd and 1st quartile, respectively. The whiskers extend from the ends of the box to the outermost data point that falls within the distances computed as follows: 3rd quartile + 1.5*(interquartile range), 1st quartile - 1.5*(interquartile range).

The Cu^{2+} initially were coordinated to two His residues for each model. Since during the MD simulations both the peptide and the Cu^{2+} are dynamics and their locations are changed, thus the Cu^{2+}-His distances are changed too during the simulations, therefore it is expected that a range of distances between Cu^{2+} and His residues will be obtained in the analysis.

Table S1: The mean conformational mean energies computed from the GBMV method.[19,20] Standard errors of the conformational mean energies were also computed.

	Conformational Energy (kcal/mol)	Standard Error (kcal/mol)
A13	-689.8	0.7
A14	-684.9	0.5
A15	-702.3	0.6
A16	-701.9	0.7
A24	-692.0	0.6
A25	-690.9	1.0
B13	-736.3	0.5
B14	-730.9	0.5
B15	-731.8	0.6
B16	-707.3	0.5
B24	-739.3	0.7
B25	-741.2	0.8

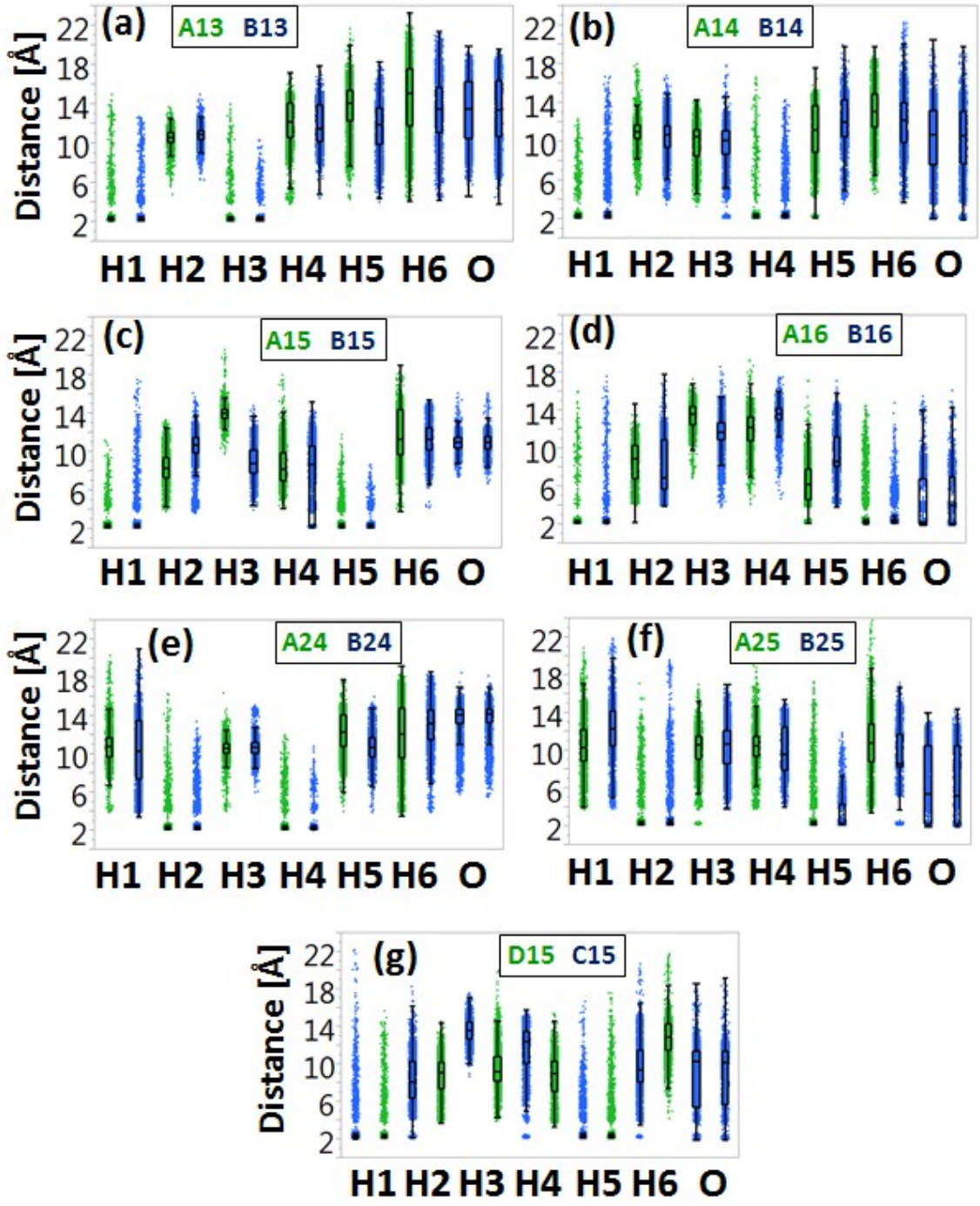


Fig. S1: The Cu^{2+}-Nε (His) atom and Cu^{2+}-O (C-terminal carboxyl group) atom distance distribution for each His residue obtained from MD simulations for models A13 and B13 (a), A14 and B14 (b), A15 and B15 (c), A16 and B16 (d), A24 and B24 (e) A25 and B25 (f) and C15 and D15 (g). The vertical lines within each box represent the median distance values.

REFERENCE

1. J. Watly, Simonovsky, E., Wieczorek, R., Barbosa, N., Miller, Y., and Kozlowski, H. Inorg Chem 2014, 53, 6675-6683.
2. Wineman-Fisher, V.; Simkovitch, R.; Shomer, S.; Gepshtein, R.; Huppert, D.; Saif, M.; Kallio, K.; Remington, S. J.; Miller, Y., *Physical Chemistry Chemical Physics* **2014,** *16*, 11196-11208.
3. Zeytuni, N.; Uebe, R.; Maes, M.; Davidov, G.; Baram, M.; Raschdorf, O.; Nadav-Tsubery, M.; Kolusheva, S.; Bitton, R.; Goobes, G.; Friedler, A.; Miller, Y.; Schueler, D.; Zarivach, R., *Plos One* **2014,** *9*.
4. Raz, Y.; Rubinov, B.; Matmor, M.; Rapaport, H.; Ashkenasy, G.; Miller, Y., *Chemical Communications* **2013,** *49*, 6561-6563.
5. Raz, Y.; Miller, Y., *Plos One* **2013,** *8*.
6. Raz, Y.; Adler, J.; Vogel, A.; Scheidt, H. A.; Haupl, T.; Abel, B.; Huster, D.; Miller, Y., *Physical Chemistry Chemical Physics* **2014,** *16*, 7710-7717.
7. Wineman-Fisher, V.; Atsmon-Raz Y.; Miller, Y. *Biomacromolecules* **2014,** 16,156-65.
8. Miller, Y.; Ma, B. Y.; Nussinov, R., *Coordination Chemistry Reviews* **2012,** *256*, 2245-2252.
9. Kale, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K., *Journal of Computational Physics* **1999,** *151*, 283-312.
10. MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M., *Journal of Physical Chemistry B* **1998,** *102*, 3586-3616.
11. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M., *Journal of Computational Chemistry* **1983,** *4*, 187-217.
12. Mahoney, M. W.; Jorgensen, W. L., *Journal of Chemical Physics* **2000,** *112*, 8910-8922.

13. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., *Journal of Chemical Physics* **1983,** *79*, 926-935.
14. Tu, K.; Tobias, D. J.; Klein, M. L., *Biophysical Journal* **1995,** *69*, 2558-2562.
15. Feller, S. E.; Zhang, Y. H.; Pastor, R. W.; Brooks, B. R., *Journal of Chemical Physics* **1995,** *103*, 4613-4621.
16. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G., *Journal of Chemical Physics* **1995,** *103*, 8577-8593.
17. Darden, T.; York, D.; Pedersen, L., *Journal of Chemical Physics* **1993,** *98*, 10089-10092.
18. Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C., *Journal of Computational Physics* **1977,** *23*, 327-341.
19. Lee, M. S.; Salsbury, F. R.; Brooks, C. L., *Journal of Chemical Physics* **2002,** *116*, 10606-10614.
20. Lee, M. S.; Feig, M.; Salsbury, F. R.; Brooks, C. L., *Journal of Computational Chemistry* **2003,** *24*, 1348-1356.
21. Kabsch, W.; Sander, C., *Biopolymers* **1983,** *22*, 2577-2637.