Доклади на Българската академия на науките
Comptes rendus de l'Académie bulgare des Sciences
Tome 77, No 2, 2024

BIOLOGY

Bioinformatics

BIODYNAMICAL IMPLICATIONS OF GRAFTING POSITION ALTERATION IN ENGINEERED MIMICS OF PARKINSON'S INHIBITOR MCoCP4

Nevena Ilieva[1]✉, Elena Lilkova[1], Peicho Petkov[2], Xubiao Peng[3], Yadi Cheng[3]

In memory of our colleague and friend Prof. Krassimir Georgiev

Received on November 22, 2023
Presented by I. Ivanov, Member of BAS, on December 19, 2023

Abstract

Proteins are currently the fastest-growing class of new therapeutic compounds but smaller proteins and peptides are generally not suitable for use as drugs. Using cyclotides – special knotted proteins stabilized by three pairs of disulphide bonds – as a transport means by grafting onto them as a scaffolding the bioactive peptides can enhance their stability, cellular uptake, and overall efficacy. Experimental methods for creating peptide aptamers are highly time- and resources-consuming. In silico approaches may speed up this process by pre-selecting the drug candidates based on certain biodynamic criteria. In this study, we probe the hypothesis about a relation between the scaffolding conformational stability in conjunction with certain plasticity upon grafting of functionally important domains and the desired biological activity of the modified through the grafting process molecules.

Key words: cyclotide scaffolding, CP4 combinatorial peptide, grafting, molecular dynamics, folding topology, Parkinson's disease

This work was supported in part by the Bulgarian National Science Fund (grant KP-06-China-10/2020) and by China-Bulgaria Intergovernmental S&T Cooperation Project at the Ministry of Science and Technology of People's Republic of China (2021.1). Computational resources were provided by BioSim HPC cluster at the Faculty of Physics, Sofia University "St. Kliment Ohridski" and by the Centre for Advanced Computing and Data Processing, supported under Grant BG05M2OP001-1.001-0003 by the Science and Education for Smart Growth Operational Programme (2014–2020).

https://doi.org/10.7546/CRABS.2024.02.06

Introduction. Cyclotides are a family of circular peptides found in various plant species [1–3]. They typically consist of 28–37 amino acid residues, which form a tail-to-head closed peptide chain with a trefoil-type knot formed by three pairs of disulphide bonds – the so-called cyclic cystine knot (CCK) motif [4]. Cyclotides exhibit a broad range of biological activities and – due to their unique topology – also exceptional conformational stability and thermal end enzymatic-degradation resistance. Coupled with their mutation tolerance and membrane penetration ability, this renders them excellent candidates for scaffoldings to display biologically active peptides thus opening a new horizon in drug design [5]. An example is the molecule MCoCP4 obtained by grafting a linear derivative of the combinatorial cyclic peptide CP4, known to reduce the toxicity of α-synuclein in yeast models [6], onto loop 6 of the inactive MCoTI-I [4]. Not only was the engineered molecule stronger in its suppression action but it would also allow for oral administration due to its resistance to proteolytic degradation and its membrane penetration potential, both inherited from the scaffolding.

In this paper, we take a systematic approach to designing potentially competitive mimics of this grafted molecule on the basis of the highly homologous trypsin inhibitor MCoTI-II by grafting the same CP4 linear derivative onto its six loops. As selection criteria, we put forward the preserved conformational stability of the scaffold after the grafting in conjunction with the conformational plasticity of the altered loop. We scrutinize the biodynamics of digital twins of the engineered mimics by means of molecular dynamics simulations. For analysis of the synthetic data we employ both GROMACS in-built tools and custom-designed approaches. Based on the obtained results, we propose several other molecular modifications with expected similar or higher inhibitory potential.

Materials and methods. ***Input structures.*** In [6], the residue D4 in loop 6 (Fig. 1) of the trypsin inhibitor MCoTI-I was experimentally substituted by the CP4 linear derivative. Our in silico study aims to explore the potential of this approach extending the procedure to the highly homologous trypsin inhibitor MCoTI-II (PDB ID 1IB9 [7]) used as a molecular matrix on the one hand, and probing grafting topology variations with grafted peptides in all of its six loops on the other.

CP4 is the combinatorial octapeptide CLATWAVG, tail-to-head connected by a peptide bond. Its linear derivative was obtained by replacing the bridge-forming cysteine with serine. In the absence of crystallographic data, the 3D structure for the subsequent in silico investigations was constructed by homology modelling, the longest matching PDB [8] motif, LATWA, being augmented with the missing three residues using UCSF Chimera software [9]. Following the approach in [10], one residue in each loop was replaced with the linearized octapeptide to generate after energy minimization the six initial structures of the sample molecule set.

Molecular dynamics (MD) simulations. All simulations were performed with the MD simulation package GROMACS 2021.1 [11], with CHARMM 36m

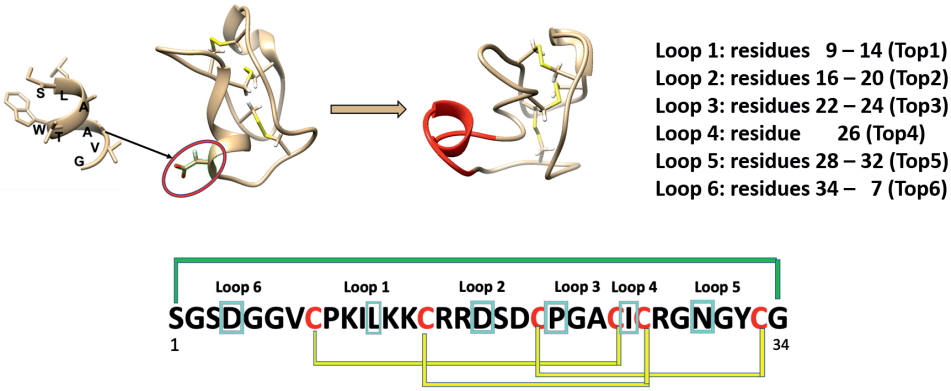


Fig. 1. (**a**) Cyclotide gtrafting scheme; (**b**) Grafting positions of CP4 derivative in MCoTI-II loops (replaced residues are boxed; disulphide bridges are shown in yellow and the head-to-tail peptide bonding – in green)

force field and the modified TIP3P water model. The package was compiled with Intel 19.0.4.2019 compiler, SIMD instructions AVX2_256 for vectorization, and with FFTW 3.3.3 used. Simulations were performed at a hybrid-architecture HPC infrastructure (Intel® Xeon® Gold 5118 CPU @ 2.30GHz, 128GB RAM & NVIDIA® Tesla® V100 GPU).

The molecules were solvated in cubic boxes, with a minimal distance of 1.2 nm to the box walls and periodic boundary conditions. Potassium and chlorine ions at a concentration of 0.15 M were added to ensure physiological conditions. After an energy minimization by the steepest descent with a maximum force tolerance of 10 kJ/(mol nm), simulated annealing was performed to set a proper temperature and to avoid the initial structure bias. After a short 100 ps position-restraint simulation to equilibrate the solvent, two nanosecond simulations were performed in NVT and NPT ensembles, with relaxation times of 0.2 ps and 0.8 ps, respectively, to stabilize the systems at a temperature of 310 K and pressure of 1 atm using Berendsen's thermo- and barostats. Finally, production simulations of 3 µs preceded by energy minimization and relaxation runs were performed in the NPT ensemble, with v-rescale temperature coupling and Parrinello–Rahman barostat, the leapfrog integrator with a time-step of 2 fs and the LINCS algorithm used to constrain the bonds between heavy atoms and hydrogens. Data was recorded every 20 ps, giving rise to 150 000 frames for each grafted system which form the basis for our analysis.[1]

Synthetic data analysis. In the MD trajectories evaluation, the standard GROMACS post-processing and analysis tools were used. For a precise quantification of the grafting-induced molecular twists and loop stiffening and their

[1]For detailed references, please consult the GROMACS Manual 2023. The scripts and config-files used are available upon request.

implications on major biodynamics trends, we employed the notion of generalized bond- and torsion angles (κ, τ) introduced within the discrete Frenet frame (DFF) formalism for protein folding and dynamics [12, 13]. Relating to four consecutive Cα atoms, these quantities form a sufficient coordinate set for reconstructing the original fractal protein geometry [14] – a feature now also being employed in the development of highly effective protein language based AI approaches for structure prediction [15]. For structure visualisation, the UCSF Chimera package [9] was used.

Results and discussion. We refer to the six different grafted molecules as topology 1 to 6 (Top1 to Top6). In all six cases, the root-mean-square deviation (RMSD) of the Cα atoms w.r.t. the initial conformation was rather small, with average values from the last third of the respective simulations between 0.3 nm and 0.6 nm, compared to 0.7 nm for the experimental-graft twin. The least volatile construct was Top1, while Top3 and Top5 were the closest to the initial conformation, and Top2 and Top6 exhibited the largest fluctuations. This RMSD data may indicate a lack of substantial global structural rearrangements but also the existence of concerted movements of (parts of) the molecules.

The individual residue flexibility substantiated through the root-mean-square fluctuations (RMSF) shown in Fig. 2 may shed some light on this matter. Thus, topologies 1, 2, 5, and 6 manifest comparable behaviour, with a more rigid core part (the scaffolding) and a relatively flexible CP4 segment, however, with some 30% smaller fluctuations along the whole chain in Top5. Top3 shows the smallest fluctuations, with less pronounced volatility of the graft but with somewhat

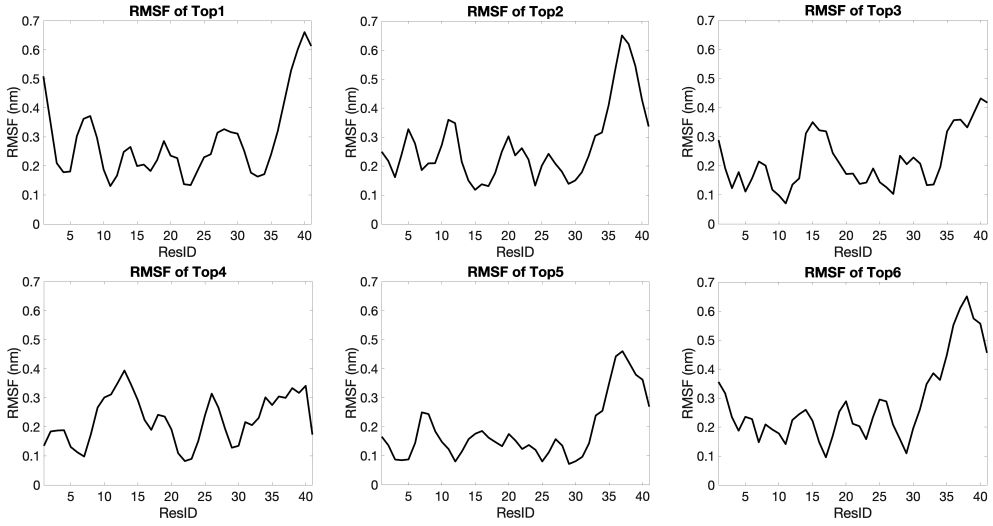


Fig. 2. The RMSF evolution of the MCoCP4 mimics in the six grafting schemes. The CP4 segment is always at the end (positions 35–42), replacements are at positions 10, 18, 22, 26, 30, and 4, respectively

greater fluctuations in loop 2 preceding the grafting domain. Top4 shows comparable fluctuations in all six loops which is indicative of substantial scaffolding-backbone deformations.

Since the molecular matrix MCoTI-II belongs to the trypsin inhibitor family, the changes upon grafting in the trypsin binding domain – residues 10–13 – are to be considered potentially impacting this feature and thus undesirable. This aspect of biodynamics can be addressed by scrutinizing the change over time of the generalized bond and torsion angles. Using as a reference the trajectory parameters for the digital twin of the experimental modification MCoCP4 based on MCoTI-I as a scaffolding, we observed increased conformation plasticity of the CP4 segment in all mimics, the smallest increase being recorded for Top1 and Top3, with Top3 maximally preserving the torsional pattern of the reference system. The trypsin binding cleft, exhibiting remarkable conformational stability in the reference system, appears to be impacted by the grafting in all six mimics though in a different way: while in Top2 and Top4 it tends to become fully disordered, in Top5 and Top6 it preserves the stability but within completely different secondary-structure elements. Top1 and Top3 are the only ones that show reasonable geometric stability and compatible angular patterns of the backbone torsion.

Targeting in the design process a stable molecular modification we focus on the stable conformations accessible for the mimics. The 2D plots relating gyration radius as a measure of the compactness of the molecule and the RMSD values discussed above are presented in Fig. 3. Top1 and Top3 show a clear tendency towards a single stable minimum, the one for Top3 being more compact and frequented and also closer in its topology to the experimental twin. The

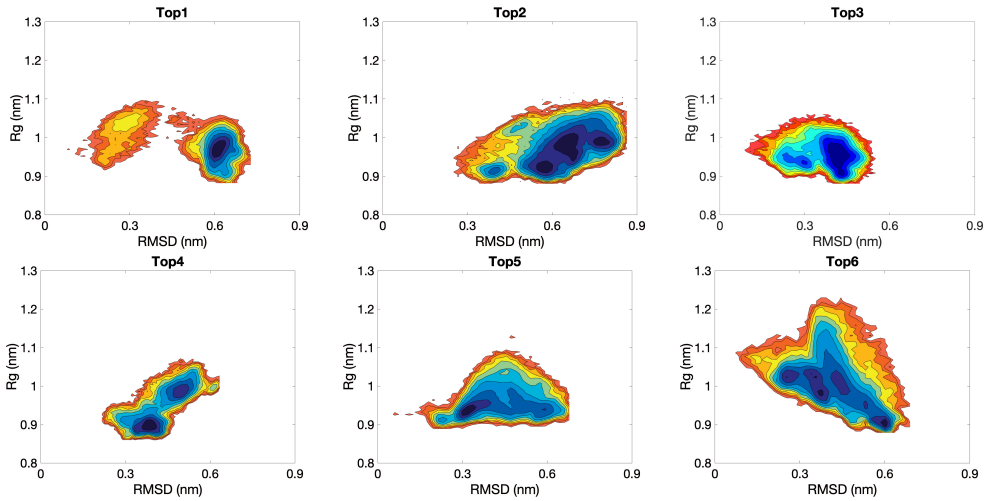


Fig. 3. Rg/RMSD plots of the MD trajectories for the six grafting schemes

single-minimum importance is further supported by the observation that in the different minima the main conformational differences are due to the grafted segment: while the scaffoldings are intertwined, the CP4 grafts adopt distinguishable conformations. As this is the active component, transitions between the minima may impact the desired bioactivity or even render the molecule inactive.

Finally, we study the differences in solvent exposure. Also here, the relative stiffness of Top1 is manifested through the narrow distributions. In both figures, three groups of states can be detected, encompassing (in decreasing order) Top6–Top2, Top3, Top4, Top5–Top1, respectively, Top1–Top2, Top3, Top6–Top4,5. The most favourable in this context appear to be Top2 and Top3, the latter showing a slightly better graft/whole-molecule SASA ratio. Moreover, Top3 is the mimic with the smallest residue-wise RMSD from the ancestor molecule MCoTI-II (Fig. 4).

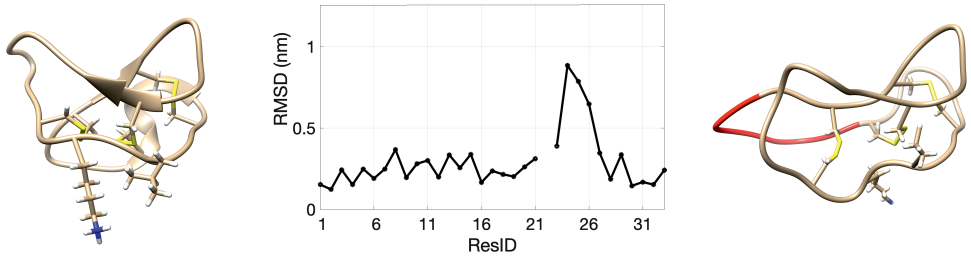


Fig. 4. (a) The ancestor molecule MCoTI-II (PDB ID 1IB); (b) Pairwise RMSD of Top3 w.r.t. MCoTI-II; (c) Final conformation of Top3 with denoted disulphide bridges, trypsin-binding-site residues, and the CP4 segment (shown in red)

Based on all of the above and also taking into account an earlier analysis of the folding dynamics of the investigated mimics, we put forward Top3 as a promising drug candidate, possibly outperforming the experimentally investigated one. The somewhat weaker performance of Top6, which shared the grafting position with the experimental reference, though on a close but still different scaffolding (MCoTI-II vs. MCoTI-I) may be attributed to the net charge difference between the two molecules in combination with the charge alteration in this topology.

Conclusions. We presented the outcome of a detailed investigation into the grafting position impact on the biodynamics of engineered cyclotide modifications with certain targeted properties. The study is the first step towards the far-reaching goal of putting together a toolkit and the corresponding protocol for in silico design of precise and highly effective pharmaceutical products for therapy and diagnostics based on appropriately modified through a grafting procedure cyclotides. The convergent selection process encourages further extension of the computational studies to encompass membrane penetration and target engagement of the constructs.

Acknowledgements. We thank L. Litov for the constructive discussions.

REFERENCES

[1] Craik D. J., N. L. Daly, T. Bond, C. Waine (1999) Plant cyclotides: a unique family of cyclic and knotted proteins that defines the cyclic cystine knot structural motif, J. Mol. Biology, **294**(5), 1327–1336.

[2] Poth A. G., M. L. Colgrave, R. Philip, B. Kerenga, N. L. Daly et al. (2011) Discovery of cyclotides in the Fabaceae plant family provides new insights into the cyclization, evolution, and distribution of circular proteins, ACS Chem. Biology, **6**(4), 345–355.

[3] Wang C. K., Q. Kaas, L. Chiche, D. J. Craik (2007) CyBase: a database of cyclic protein sequences and structures, with applications in protein discovery and engineering, Nucleic Acids Res., **36**, 206–210.

[4] Schwarz E. (2017) Cystine knot growth factors and their functionally versatile proregions, Biol. Chem., **398**(12), 1295–1308.

[5] Senthilkumar B., R. Rajasekaran (2017) Analysis of the structural stability among cyclotide members through cystine knot fold that underpins its potential use as a drug scaffold, Int. J. Pep. Res. Ther., **23**(1), 1–11.

[6] Kritzer J. A., S. Hamamichi, J. M. McCaffery, S. Santagata, T. A. Naumann et al. (2009) Rapid selection of cyclic peptides that reduce α-synuclein toxicity in yeast and animal models, Nature Chem. Biol., **5**(9), 655–663.

[7] Felizmenio-Quimio M. E., N. L. Daly, D. J. Craik (2001) Circular proteins in plants: solution structure of a novel macrocyclic trypsin inhibitor from Momordica cochinchinensis, J. Biol. Chem., **276**(25), 22875–22882.

[8] Berman H. M., J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat et al. (2000) The protein data bank, Nucl. Acids Res., **28**(1), 235–242.

[9] Pettersen E. F., T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt et al. (2004) UCSF Chimera – a visualization system for exploratory research and analysis, J. Comp. Chem., **25**(13), 1605–1612.

[10] Jagadish K., A. Gould, R. Borra, S. Majumder, Z. Mushtaq et al. (2015) Recombinant expression and phenotypic screening of a bioactive cyclotide against α-synuclein-induced cytotoxicity in baker's yeast, Ang. Chem., **127**(29), 8510–8514.

[11] Abraham M., T. Murtola, R. Schulz, S. Páll, J. C. Smith et al. (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX, **1**, 19–25.

[12] Krokhotin A., A. Niemi, X. Peng (2012) Soliton concepts and protein structure, Phys. Rev. **E85**, 031906.

[13] Dai J., A. J. Niemi, J. He, A. Sieradzan, N. Ilieva (2016) Bloch spin waves and emergent structure in protein folding with HIV envelope glycoprotein as an example, Phys. Rev. **E93**(3), 032409.

[14] Hinsen K., S. Hu, G. R. Kneller, A. J. Niemi (2013) A comparison of reduced coordinate sets for describing protein structure, J. Chem. Phys., **139**, 124115.

[15] Chowdhury R., N. Bouatta, S. Biswas, C. Floristean, A. Kharkar et al. (2022) Single-sequence protein structure prediction using a language model and deep learning, Nature Biotechnol., **40**, 1617–1623.

[16] Jones P. M., A. M. George (2016) Computational analysis of the MCoTI-II plant defence knottin reveals a novel intermediate conformation that facilitates trypsin binding, Nature Sci. Rep., **6**, 23174.

[1]*Institute of Information and Communication Technologies,*
Bulgarian Academy of Sciences, Akad. G. Bonchev St, Bl. 25A, 1113 Sofia, Bulgaria
e-mails: nevena.ilieva@iict.bas.bg, elena.lilkova@iict.bas.bg

[2]*Faculty of Physics, Sofia University "St. Kl. Ohridski",*
5 J. Bourchier Blvd, 1164 Sofia, Bulgaria
e-mail: peicho@phys.uni-sofia.bg

[3]*Center for Quantum Technology Research and Key Laboratory of Advanced Optoelectronic Quantum Architecture and Measurements, School of Physics, Beijing Institute of Technology, Beijing 100081, China*
e-mail: xubiaopeng@bit.edu.cn, yadicheng1@gmail.com