

## Fast and Reliable Folding of Metalloproteins

Combining reservoir replica exchange simulations (R-REMD) with TIGER2hs algorithm and flat bottom constraints to fold the atypical zinc finger domain of the transcription factor BCL11B

L. Schulig<sup>1</sup>, H. Forke<sup>2</sup>, P. Grabarczyk<sup>2</sup>,  
C. A. Schmidt<sup>2</sup>, A. Link<sup>1\*</sup>,

<sup>1</sup>Institute of Pharmacy, University of Greifswald,

<sup>2</sup>Internal Medicine C, University of Greifswald

### In Short

- Transcription factor BCL11B
- Atypical zinc finger domain
- Protein folding using enhanced sampling methods
- Reservoir replica exchange simulations with TIGER2hs algorithm

**Introduction** Transcription factors play a crucial role in regulating biological processes like cell growth, differentiation, organ development and cellular signaling [1]. Within this group proteins equipped with zinc finger motifs (ZFs) represent the largest family of sequence-specific DNA binding transcription regulators.

Studies have proven the relevance of BCL11B for a variety of tissues and organs. Loss of BCL11B in mice knockout models showed severe developmental abnormalities of central nervous system, T cells, skin, teeth, and mammary glands [2,3]. Moreover, a very intriguing phenomenon was revealed upon inducible BCL11B knockout in mouse thymocytes and fully differentiated T cells. These T-lineage committed cells re-differentiated into NK cell-like chimeras with very strong proliferating potential and cytotoxic properties providing efficient immune defense against tumor cells *in vivo*.

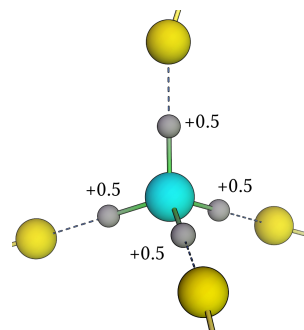
The similar role of BCL11B in humans had been confirmed by various pathologies observed mostly but not only in T cells and CNS. Multiple diseases like some types of leukemia, neurodegenerative or inflammatory disorders result from deregulated BCL11B expression or genetic aberrations involving the BCL11B gene [4].

BCL11B belongs to the Krüppel-like zinc finger sub-family of proteins and encodes six classical CCHH zinc finger (ZF) domains mediating sequence-specific DNA binding. The CCHC zinc finger is located at the N-terminus of the protein and structurally resembles CCHC motifs of FOG-family proteins [5,6]. Its unique sequence and importance for proper BCL11B function makes it an ideal target for protein inhibition studies. The ability to block

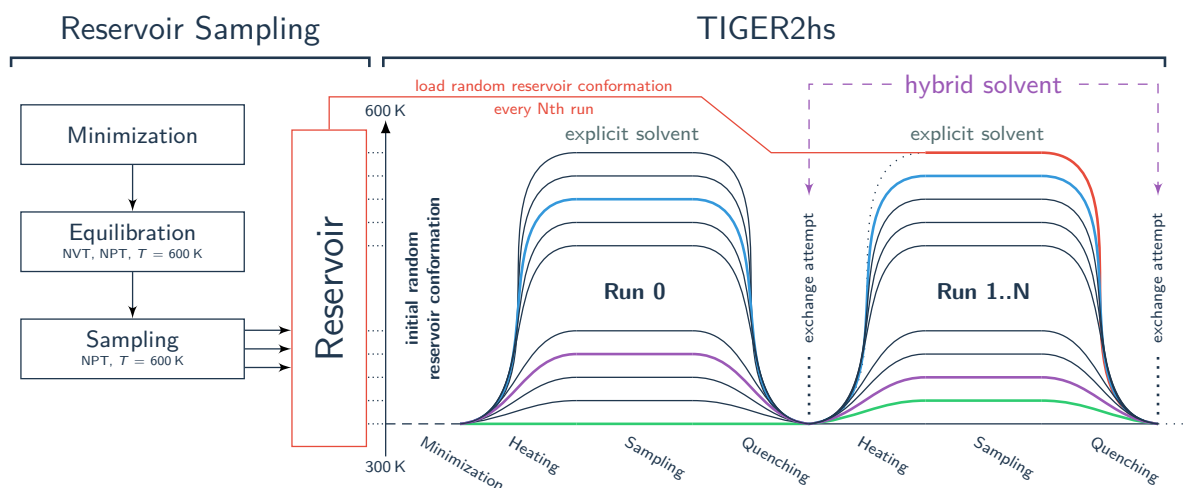
BCL11B protein in turn might possess multiple applications including therapy of BCL11B-dependent malignancies or reprogramming of T-cells into professional killer cells for cellular therapies. This however must be preceded by studies on the CCHC structure which will allow characterizing the interaction interface.

**TIGER2hs** Conventional replica exchange methods are too computationally demanding because a large number of replicas with small temperature differences are necessary for meaningful exchange rates, which has an dramatic effect on the convergence[7,8]. Another problem is the significantly larger number of water than protein atoms, leading to fluctuations and noise in the total energy compared to the protein alone. TIGER2hs solves this problem by cooling all replicas to the baseline temperature after a sampling period, calculating the energy with the first two water shells around the protein in implicit solvent (hybrid solvent) and heat them up again after the exchange attempt and ranking (figure 2). The creation of secondary structure elements is reliable and more accurate than implicit solvent models only.

**Cationic Dummy Atom Model** In molecular dynamics simulations based on classical force fields, metal ions are treated as a single point charge. Coordinative bonds are therefore described by van der Waals (vdW) and electrostatic interactions only. This leads to stability issues, abnormal coordination numbers or unphysical behavior. Using a cationic dummy atom model [9], we are able to distribute the charge around a center particle in the form of small dummy atoms within the vdW-radius of a normal  $Zn^{2+}$  ion.



**Figure 1:** Four cystein sulfur atoms bound to Cationic Dummy Atom Model



**Figure 2:** Improved version of the TIGER2hs algorithm extended by a precalculated reservoir of randomly sampled conformations.

**R-REMD** The reservoir replica exchange approach was initially developed to improve convergence of the simulations [10] but in our case it also ensures random conformations without cysteins bonded to the zinc after every Nth iteration. Even with distributed charges in the cationic dummy atom model (figure 1), it is very unlikely that a recoordination occurs within the simulation time range. The reservoir conformations are precalculated at 600 K without any charged interaction to the zinc ion for a randomly sampled amino acid chain. Initial conformations for each replica and the highest temperature replica at every Nth step are replaced by a randomly picked structure from this reservoir (figure 2).

**Flat bottom constraints** An additional, linearly-scaled force acts on the cystein sulfur atoms when they are not within a specified distance to the zinc ion. This facilitates weak boundaries and guidance for creating the metal complex.

**Objective** The aim of this project is to establish and validate the novel R-TIGER2hs algorithm as a general method for *in silico* folding of metal depend proteins. Afterwards it will be used to predict possible tertiary structures of the atypical zinc finger domain of the transcription factor BCL11B. These are important for simulating and understanding the dimerization process, laying the foundation for drug discovery and antibody design.

- [2] K. Kyrylkova et al., *Plos One* **7**, e37670 (2012). doi:10.1371/journal.pone.0037670
- [3] Y. Wakabayashi et al., *Nat Immunol* **4**, 533–539 (2003). doi:10.1038/ni927
- [4] Y. Wakabayashi et al., *Biochem Biophys Res Commun* **301**, 598–603 (2003). doi: 10.1016/S0006-291X(02)03069-3
- [5] P. Grabarczyk et al., *Mol Cell Biol* **38**, e00368-17 (2018). doi:10.1128/MCB.00368-17
- [6] J.M. Matthews et al., *Eur J Biochem* **267**, 1030–1038 (2000). doi:10.1046/j.1432-1327.2000.01095.x
- [7] D.J. Earl, M.W. Deem, *Phys Chem Chem Phys* **7**, 3910–3916 (2005). doi: 10.1039/B509983H
- [8] M. Kulke et al., *J Phys Chem B* **122**, 7295–7307 (2018). doi:10.1021/acs.jpccb.8b05178
- [9] Y. P. Pang, *Proteins* **45**, 183–189 (2001). doi: 10.1002/prot.1138
- [10] A.E. Roitberg et al., *J Phys Chem B* **111**, 2415–2418 (2007). doi:10.1021/jp068335b

### Project Partners

Prof. W. Langel, N. Geist, M. Kulke, Biophysical Chemistry Group, University of Greifswald

### More Information

- [1] C. Stienne et al., *Immunity* **45**, 774–787 (2016). doi:10.1016/j.immuni.2016.09.010