

## Construction of Brn2 Protein in Complex with MORE DNA Through *In Silico* Methods

Ivan Evangelista do Vale Coelho<sup>a</sup> (PG), Denise Costa Arruda<sup>b</sup> (PQ), Alex G. Taranto<sup>a</sup> (PQ)

<sup>a</sup> Federal University of São João del-Rei, Campus Centro-Oeste Dona Lindu, Divinópolis/MG, Brazil.

<sup>b</sup> University of Mogi das Cruzes, Mogi das Cruzes/SP, Brazil  
ivanglista2009@yahoo.com.br

Keywords: Melanoma, Transcription Factor, N-Oct-3, Molecular Dynamics Simulation

### INTRODUCTION

The incidence of non-melanoma (basal cell and squamous cell carcinoma) and melanoma skin cancers has increased in last decades. In three diagnosed cancers, one is skin cancer. More rare and lethal, melanoma is the most aggressive type of skin cancer. Currently, 132000 cases occur worldwide each year.<sup>1</sup> However, if detected early, the cure chances are more than 90%. Metastatic melanoma cases have a worse prognosis, which in most cases has no cure and have a limited number of therapeutic options.<sup>2</sup> Therefore, prevention campaigns and early diagnosis are very important. Typically, melanoma arises in areas of the body most exposed to solar radiation, especially people with pale skin.<sup>3</sup> In recent years, some new drugs have improved the survival of patients with melanoma. However, this increase in survival is modest and most patients with metastatic melanoma are not cured.<sup>2</sup> Thus, studies should be conducted to search for new molecular targets and more effective treatment options for melanoma skin cancer.

In this context, Brn2 transcription factor has been studied due its relationship with melanoma development. High levels of this protein expression, compared to the low levels found in differentiated melanocytes, have been found in melanoma cell lines.<sup>4,5,6,7,8</sup> The Brn2 protein, codified by POU3F2 gene and known as N-Oct-3 when complexed with DNA target site, is a member of transcription factor family, which has conserved POU domain.<sup>9</sup> Melanocyte precursors also express Brn2, but express lower levels with the differentiation of these cells. This protein can coordinate the normal development of melanocytic lineage or can reactivate signals to

an abnormal growth as in malignant melanoma.<sup>10</sup> Moreover, there is evidence that BRN2 overexpression is regulated by the major signaling pathways in melanoma: Wnt/b-catenin, MAPK/BRAF e PI3K/AKT.<sup>8,11,12</sup>

The Brn2 structure can interact with distinct types of DNA targets: the more palindromic Oct recognition element (MORE), palindromic Oct recognition element (PORE) and N-Oct-3 recognition element (NORE) as a monomer but also as a dimer.<sup>13</sup> Similar to POU proteins, Brn2 can interact with several DNA sequences, with different orientations and spacing between its POU-specific subdomain (POUs) and POU-homeodomain (POUh) subdomains because of subdomains are bonded by flexible linker. This protein, constituted of 443 amino acids (47 kDa), has both transactivation N-terminal domain (local of interactions with co-regulatory proteins) and DNA-binding C-terminal domain.<sup>14</sup> The conserved DNA-binding domain consists of N-terminal POU of approximately 75 amino acids and C-terminal POUh of 60 amino acids joined by a less conserved linker.<sup>9</sup> The POU domain-DNA interaction occurs primarily on 5'-ATGCAAAT-3' octameric sequence. Both POU and POUh subdomains form a helix-turn-helix motif using the second and third helices of each subdomain.<sup>15</sup> However, neither the Brn2 three-dimensional structure nor its POU domain are available in the Protein Data Bank (PDB).<sup>16</sup> Only part of the structure (POU domain) of several POU proteins can be found in PDB, and no complete structure is available. The Brn2 DNA-binding domain was constructed through homology modeling using as template the Oct-1 protein (PDB code: 1OCT).<sup>17</sup> However, the construction of full-

length Brn2 model and molecular dynamics (MD) simulation of this protein in complex with MORE DNA provide knowledge of its structure and function. These studies can contribute to the development of selective inhibitors against this protein or DNA target site.

In this work, we construct the full-length three-dimensional structure of Brn2 bound to MORE half-site (5'-ATGCATGAGGA-3'). The MD simulation of Brn2-MORE complex is in progress.

## METHODS

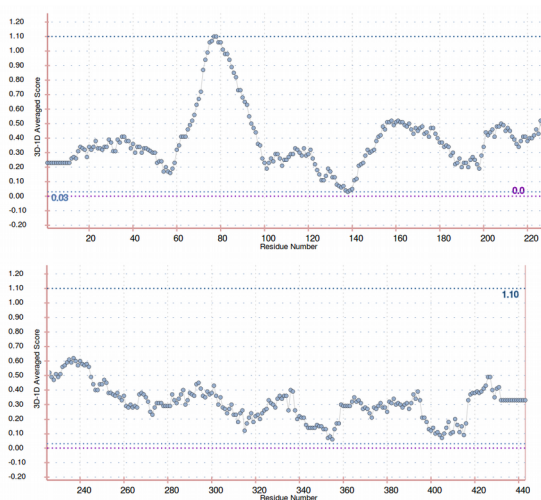
The primary sequence of Brn2 was retrieved from NCBI's protein database,<sup>18</sup> access code NP\_005595.2.<sup>19</sup> Currently, there is no template available in PDB for full Brn2 protein. Thus, this protein was built using PHYRE2 program.<sup>20</sup> This built model was improved using Swiss-Model program,<sup>21</sup> which built a new brn2 model using the previous model generated by PHYRE2 and Oct-6 (POU3F1) POU domain available in PDB (2XSD) as templates. The model was evaluated to structural quality using PROCHECK 3.5.4,<sup>22</sup> Verify 3D,<sup>23,24</sup> ANOLEA<sup>25</sup> and structural features obtained from circular dichroism spectra of full-length Brn2 protein.<sup>14</sup> The DNA-target structure which has the 5'-ATGCATGAGGA-3' sequence was retrieved from PDB (2XSD). This complex was submitted to 20000 steps of energy minimization under conjugated gradient using NAMD program<sup>26</sup> and the CHARMM36 all-atom force field.<sup>27,28,29,30</sup> The explicit solvent model with water box (TIP3P water model)<sup>31,32</sup> and periodic boundary conditions were used. The protein-DNA complex was immersed in a water box and solvent molecules closer than 2.4 Å to the solute were eliminated. The dimensions of box sides were 79, 90 and 104 Å whereas the water box had at least 12 Å of extension between any protein-DNA atom and the edge of the box. Counter ions Na<sup>+</sup> were added to neutralize the system before simulation. Electrostatic interactions were calculated using the Particle Mesh Ewald method<sup>33</sup> with real-space cutoff of 12 Å and 1 Å between grid points. The Lennard-Jones interactions were included with force switching from 10 Å to 12 Å. The list of

nonbonded atoms was kept for interatomic distances of 14 Å.

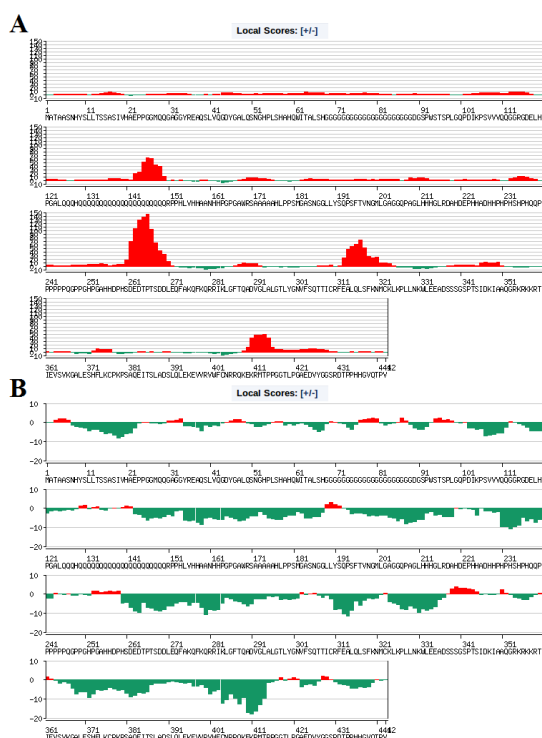
## RESULTS AND DISCUSSION

The full three-dimensional structure of the Brn2 protein was obtained by comparative modeling method. This protein was bound to an 11 base pair oligonucleotide containing the half-site of the MORE motif (ATGCAT) and optimized through energy minimization. The stereochemical quality of the Brn2 model, was evaluated using the PROCHECK 3.5.4 software after energy minimization. The model showed 76.5% of its amino acid residues in favourable regions, 17.9% in allowed regions, 2.7% in generously allowed regions and 3.0% in disallowed regions (except glycine and proline residues). Moreover, the POU domain, which interacts with DNA, has 92.7% of its amino acid residues in favourable regions and 5.8% in allowed regions. This region (both POU and POUh subdomains) has 95.31% of sequence identity with the same Oct-6 region available in PDB which was also used for model building. High sequence similarity level (greater than 30%) confers reliability to built model.<sup>34</sup> The model analysis through Verify 3D showed that 85.78% of the residues had an averaged 3D-1D score  $\geq 0.2$  (threshold value is 80%)<sup>23,24</sup> (figure 1). The model showed compatibility of its atomic model (3D) with its own amino acid sequence (1D) when comparing the results to good structures. The figure 2 shows the energy estimation of Brn2 model through non-local atomic interaction energy assessment using ANOLEA before and after energy minimization. The model quality was improved after energy minimization. As can be seen, the number and energy value of high-energy amino acid residues decrease after the process. The POU domain of this protein is constituted mainly of helices and transactivation domain of loop regions. As shown in figure 3, the model has 9  $\alpha$ -helices without  $\beta$ -sheet. The POU subdomain consists of 4  $\alpha$ -helices and POUh of 3  $\alpha$ -helices. Both subdomains have helix-turn-helix motif using the second and third helices of each subdomain, similar to other POU proteins. The protein remaining is mainly constituted of loop regions and two short helices (figure 3). The Brn2 structure was compared with structural features data obtained from circular dichroism

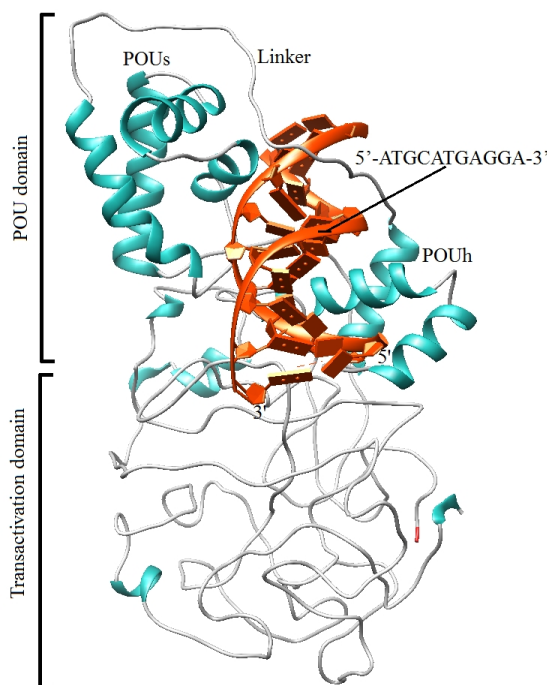
spectra of full-length Brn2 reported by Cabos-Siguier and collaborators.<sup>14</sup> The structure of Brn2 model was similar to circular dichroism data which suggested the presence of  $\alpha$ -helical secondary structures. The location of tryptophan residues was also in agreement with the circular dichroism spectra mentioned above. The side chains of the three tryptophan residues located in the transactivation domain were more exposed to the solvent than those two tryptophan residues found in the POU subdomains, which was buried inside the three-dimensional structure of the protein. Thus, the model showed good quality for further MD simulation in complex with MORE half-site which is in progress. This studies can contribute to describe the structural properties of full Brn2 protein and the mechanism of its interaction with specific DNA target site.



**Figure 1.** Analysis of Brn2 model using Verify 3D. The model has 85.78% of the residues with averaged 3D-1D score  $\geq 0.2$  (threshold value is 80%).



**Figure 2.** Energy estimation of Brn2 model using ANOLEA before (A) and after (B) energy minimization. The plot shows a detailed energy analysis along the sequence.



**Figure 3.** Three-dimensional structure of full-length Brn2 model in complex with MORE half-site. The recognition helices of Brn2 POUh and Brn2 POUl are inserted within the major groove of the DNA (red). The model is

constituted mainly of helices (blue) and loop (white) regions.

## CONCLUSIONS

The model showed good quality verified through evaluation by PROCHECK 3.5.4, Verify 3D, ANOLEA and structural features obtained from circular dichroism spectra of Brn2 protein. The POU domain of Brn2 model shared structural similarity to other members of transcription factor family. This Brn2 domain is also very similar to Oct-6 POU domain available in PDB, which the regions share high sequence similarity. The MD simulation of Brn2-DNA target site is in progress. This simulation will allow to refine this protein structure, and understand its specific interactions with MORE DNA. Knowledge about these interactions will also provide information to contribute for the development of specific inhibitors against Brn2 or DNA target site.

## ACKNOWLEDGMENTS

The authors are grateful for the support given from the CAPES, PPGBiotech/UFSJ, FAPEMIG and FAEP.

<sup>1</sup> World Health Organization (WHO), 2015. Skin Cancers. Retrieved from <http://www.who.int/uv/faq/skincancer/en/>.

<sup>2</sup> C. Karimkhani, R. Gonzalez, R. P. Dellavalle, *Am. J. Clin. Dermatol.*, 15, 323 (2014).

<sup>3</sup> U. Leiter and C. Garbe, *Adv. Exp. Med. Biol.*, 624, 89, (2008).

<sup>4</sup> R. A. Sturm, B. J. O'Sullivan, J. A. Thomson, N. Jamshidi, J. Pedley, P. G. Parsons, *Pigment Cell Res.*, 7, 235, (1994).

<sup>5</sup> J. A. Thomson, K. Murphy, E. Baker, G. R. Sutherland, P. G. Parsons, R. A. Sturm, F. Thomson, *Oncogene*, 11, 691, (1995).

<sup>6</sup> T. Eisen, D. J. Easty, D. C. Bennett, C. R. Goding, *Oncogene*, 11, 2157, (1995).

<sup>7</sup> A. Flammiger, R. Besch, A. L. Cook, T. Maier, R. A. Sturm, C. Berking, *J. Invest. Dermatol.*, 129, 945, (2009).

<sup>8</sup> J. Goodall, C. Wellbrock, T. J. Dexter, K. Roberts, R. Marais, C. R. Goding, *Mol. Cell Biol.*, 24, 2923, (2004).

<sup>9</sup> A. K. Ryan and M. G. Rosenfeld, *Genes Dev.*, 11, 1207, (1997).

<sup>10</sup> R. Besch and C. Berking, *Eur. J. Cell Biol.*, 93, 55, (2014).

<sup>11</sup> J. Goodall, S. Martinozzi, T. J. Dexter, D. Champeval, S. Carreira, L. Larue, C. R. Goding, *Mol. Cell Biol.*, 24, 2915, (2004).

<sup>12</sup> E. Bonvin, P. Falletta, H. Shaw, V. Delmas, C. R. Goding, *Mol. Cell Biol.*, 32, 4674, (2012).

<sup>13</sup> L. Nieto, G. Joseph, A. Stella, P. Henri, O. Burlet-Schiltz, B. Monsarrat, E. Clottes, M. Erard, *J. Mol. Biol.*, 370, 687 (2007).

<sup>14</sup> B. Cabos-Siguier, A. L. Steunou, G. Joseph, R. Alazard, M. Ducoux-Petit, L. Nieto, B. Monsarrat, M. Erard, E. Clottes. *Protein Express. Purif.*, 64, 39, (2009).

<sup>15</sup> A. L. Cook and R. A. Sturm, *Pigment Cell Melanoma Res.*, 21, 611, (2008).

<sup>16</sup> H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, *Nucl. Acids Res.*, 28, 235, (2000). Available in [www.rcsb.org](http://www.rcsb.org).

<sup>17</sup> S. Millevoi, L. Thion, G. Joseph, C. Vossen, L. Ghisolfi-Nieto, M. Erard, *Eur. J. Biochem.*, 268, 781 (2001).

<sup>18</sup> K. D. Pruitt, G. R. Brown, S. M. Hiatt, F. Thibaud-Nissen, A. Astashyn, O. Ermolaeva, C. M. Farrell, J. Hart, M. J. Landrum, K. M. McGarvey, M. R. Murphy, N. A. O'Leary, S. Pujar, B. Rajput, S. H. Rangwala, L. D. Riddick, A. Shkeda, H. Sun, P. Tamez, R. E. Tully, C. Wallin, D. Webb, J. Weber, W. Wu, M. Dicuccio, P. Kitts, D. R. Maglott, T. D. Murphy, J. M. Ostell, *Nucleic Acids Res.*, 42, 756, (2014).

<sup>19</sup> C. L. Avery, C. M. Sitlani, D. E. Arking, D. K. Arnett, J. C. Bis, E. Boerwinkle, B. M. Buckley, Y. D. Ida Chen, A. J. de Craen, M.

- Eijgelsheim, D. Enquobahrie, D. S. Evans, I. Ford, M. E. Garcia, V. Gudnason, T. B. Harris, S. R. Heckbert, H. Hochner, A. Hofman, W. C. Hsueh, A. Isaacs, J. W. Jukema, P. Knekt, J. A. Kors, B. P. Krijthe, K. Kristiansson, M. Laaksonen, Y. Liu, X. Li, P. W. Macfarlane, C. Newton-Cheh, M. S. Nieminen, B. A. Oostra, G. M. Peloso, K. Porthan, K. Rice, F. F. Rivadeneira, J. I. Rotter, V. Salomaa, N. Sattar, D. S. Siscovick, P. E. Slagboom, A. V. Smith, N. Sotoodehnia, D. J. Stott, B. H. Stricker, T. Stürmer, S. Trompet, A. G. Uitterlinden, C. van Duijn, R. G. Westendorp, J. C. Witteman, E. A. Whitsel, B. M. Psaty, *Pharmacogenomics J.*, 14, 6, (2014).
- <sup>20</sup> L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass, M. J. Sternberg, *Nat. Protoc.*, 10, 845, (2015).
- <sup>21</sup> M. Biasini, S. Bienert, A. Waterhouse, K. Arnold, G. Studer, T. Schmidt, F. Kiefer, T. G. Cassarino, M. Bertoni, L. Bordoli, T. Schwede, *Nucleic Acids Res.*, 42, 252, (2014).
- <sup>22</sup> R. A. Laskowski, M. W. MacArthur, D. S. Moss, J. M. Thornton, *J. App. Cryst.*, 26, 283, (1993).
- <sup>23</sup> J. U. Bowie, R. Lüthy, D. Eisenberg, *Science*, 253, 164, (1991).
- <sup>24</sup> R. Lüthy, J. U. Bowie, D. Eisenberg, *Nature*, 356, 83, (1992).
- <sup>25</sup> F. Melo and E. Feytmans, *J. Mol. Biol.*, 277, 1141, (1998).
- <sup>26</sup> J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, K. Schulten, *J. Comput. Chem.*, 26, 1781 (2005).
- <sup>27</sup> R. B. Best, X. Zhu, J. Shim, P. E. Lopes, J. Mittal, M. Feig, A. D. MacKerell Jr., *J. Chem. Theory Comput.*, 8, 3257, (2012).
- <sup>28</sup> A. D. MacKerell Jr., M. Feig, C. L. III. Brooks, *J. Am. Chem. Soc.*, 126, 698, (2004).
- <sup>29</sup> K. Hart, N. Foloppe, C. M. Baker, E. J. Denning, L. Nilsson, A. D. MacKerell Jr., *J. Chem. Theory Comput.*, 8, 348, (2012).
- <sup>30</sup> A. D. MacKerell Jr. and N. K. Banavali, *J. Comput. Chem.*, 21, 105, (2000).
- <sup>31</sup> M. W. Mahoney and W. L. Jorgensen, *J. Chem. Phys.*, 112, 8910, (2000).
- <sup>32</sup> W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.*, 79, 926, (1983).
- <sup>33</sup> T. Darden, D. York, L. Pedersen, *J. Chem. Phys.*, 98, 10089, (1993).
- <sup>34</sup> L. Bordoli, F. Kiefer, K. Arnold, P. Benkert, J. Battey, T. Schwede, *Nat. Protoc.*, 4, 1, (2009).