

## ***In silico* Analysis and Site-directed Mutagenesis of Promoters**

Salah Boudjadi<sup>1,\*</sup> and Jean-Francois Beaulieu<sup>2,\*</sup>

<sup>1</sup>Cancer Molecular Pathology Section, Laboratory of Pathology, National Cancer Institute, National Institutes of Health, Bethesda, USA; <sup>2</sup>Laboratory of Intestinal Physiopathology, Department of Anatomy and Cell Biology, Faculty of Medicine and Health Sciences, Université de Sherbrooke, Sherbrooke, Canada

\*For correspondence: [Jean-Francois.Beaulieu@USherbrooke.ca](mailto:Jean-Francois.Beaulieu@USherbrooke.ca); [Salah.Boudjadi@nih.gov](mailto:Salah.Boudjadi@nih.gov)

**[Abstract]** In normal as in cancerous cells, gene expression is tightly regulated by transcription factors, which are responsible for up- or down-regulation of thousands of targets involved in different cell processes. Transcription factors can directly regulate the expression of genes by binding to specific DNA sequences known as response elements. Identification of these response elements is important to characterize targets of transcription factors in order to understand their contribution to gene regulation. Here, we describe *in silico* analysis coupled to selected mutagenesis and promoter gene reporter assay procedures to identify and analyze response elements in the proximal promoter sequence of genes.

**Keywords:** Site-directed mutagenesis, *in silico* analysis, Gene expression, Promoter gene-reporter assay, Response elements, Transcription factors

**[Background]** The impact of a transcription factor on global gene expression can be studied through its knockdown using shRNA or CRISPR-Cas9 methods followed by microarray analysis which can provide significant data on hundreds of dysregulated genes. This analysis, although very useful, lacks information about the direct control of the dysregulated genes. To further investigate these genes, *in silico* analysis using bioinformatics provides additional information to identify direct targets of the transcription factor studied. Additionally, the functionality of these response elements can be analyzed using site-directed mutagenesis and promoter-gene-reporter assays. We have shown recently that the oncogenic transcription factor MYC (Cellular Myelocytomatosis Oncogene) controls the expression of the *ITGA1* (integrin alpha 1 subunit) gene in colorectal cancer, and that their expression correlates in 72% of colorectal tumors. This protocol describes a procedure for the analysis of the *ITGA1* promoter and the identification of response elements for the MYC oncogene. The latter is known to be a member of the MYC/MAX/MAD network. Interactions between these factors lead to gene activation or repression depending on upstream signalization and cell condition.

### **Materials and Reagents**

1. 1.5 ml tubes
2. 10 cm dish
3. 12-well plates (Corning, Falcon®, catalog number: 353043)

4. White 96-well assay plate (Corning, catalog number: 3912)
5. Ice
6. HEK293T cells (ATCC, catalog number: CRL-3216) at low passage
7. DH5 $\alpha$  competent cells
8. Primers (Sequences of the oligonucleotide primers used for the site directed mutagenesis):  
 Forward 5'-CGACTTCACGGTGAATTTGGACAATCCGCAGGGGATGGAAGG-3'  
 Reverse 5'-CCTTCCATCCCCTGCGGATTGTCCAAATTCACCGTGAAGTCG-3'
9. The promoter of interest, as for example here, the *ITGA1* proximal promoter sequence inserted in the pLightSwitch\_prom plasmid vector (SwitchGear Genomics, catalog number: S706788)
10. Plasmid constructs carrying the activator or repressor to be characterized, as for example here: pcDNA-Empty vector, pcDNA-MYC and pCMV-MAD, as described in Ni *et al.*, 2005
11. The pGL4.13 plasmid (luc2/SV40) as a control of transfection efficiency
12. GeneArt Site-Directed Mutagenesis Kit (Thermo Fisher Scientific, Invitrogen™, catalog number: A13282)
13. DNase and RNase free water
14. Distilled water
15. 0.5 M EDTA, pH 8
16. SOC medium
17. LB agar plates
18. Trypan blue
19. Effectene transfection reagent (QIAGEN, catalog number: 301425)
20. PBS
21. DMEM medium (Thermo Fisher Scientific, Gibco™, catalog number: 11995073)
22. Fetal bovine serum (FBS) (BOLLE COMMUNICATION, WISENT, catalog number: 080-150)
23. GlutaMAX (Thermo Fisher Scientific, Gibco™, catalog number: 35050061)
24. HEPES (BOLLE COMMUNICATION, WISENT, catalog number: 330-050-EL)
25. Dual Luciferase Reporter Assay System (Promega, catalog number: E1980)  
 Luciferase Assay substrate  
 Luciferase Assay buffer
26. DMEM culture medium for HEK293T cells (see Recipes)
27. LAR II solution (see Recipes)
28. 1x Stop & Glo (see Recipes)

### **Equipment**

1. Orion Microplate Luminometer (Berthold, Bad Wildbad, Germany)
2. MyCycler Personal Thermal Cycler for PCR reactions (Bio-Rad Laboratories)
3. Water bath
4. Cell culture incubator

## **Software**

1. MatInspector web based software. Genomatix (Munich, Germany)  
[https://www.genomatix.de/online\\_help/help\\_matinspector/matinspector\\_help.html](https://www.genomatix.de/online_help/help_matinspector/matinspector_help.html)
2. Blast Global Align  
[https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE\\_TYPE=BlastSearch&PROG\\_DEF=blastn&BLAST\\_PROG\\_DEF=blastn&BLAST\\_SPEC=GlobalAln&LINK\\_LOC=BlastHomeLink](https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch&PROG_DEF=blastn&BLAST_PROG_DEF=blastn&BLAST_SPEC=GlobalAln&LINK_LOC=BlastHomeLink)

## **Procedure**

1. Promoter analysis
  - a. The proximal promoter sequence of the promoter of interest, the *ITGA1* gene, can be downloaded from the reference human genome or from the switchdb website:  
[http://switchdb.switchgeargenomics.com/productinfo/id\\_706788/](http://switchdb.switchgeargenomics.com/productinfo/id_706788/)
  - b. Genomatix website: upload the promoter sequence and use the default parameters or use the promoter search.
  - c. Guideline detailed in:  
[https://www.genomatix.de/online\\_help/help\\_matinspector/matinspector\\_help.html#eldo\\_promoters](https://www.genomatix.de/online_help/help_matinspector/matinspector_help.html#eldo_promoters)
  - d. The identified response elements are compared to the consensus sequence (which binds the transcription factor). For instance, for MYC, possible response elements are compared to the E-box element CACGTG. In addition to the results obtained with the software, visual analysis of the promoter sequence is highly recommended.
2. Site directed mutagenesis  
 Design primers manually and change nucleotides in order to completely modify the sequence of the identified response elements (avoid creating a response element for another transcription factor). The mutations should be centrally located in both the forward and reverse primers. The primers should be complementary and 35 to 45 nucleotides in length to insure proper annealing.
  - a. Prepare the PCR mix (excepting primers, all reagents are included in the GeneArt Site-Directed Mutagenesis Kit)

10x AccuPrime™ Pfx reaction mix	5 µl	1x
10x enhancer	5 µl	1x
Primer mix (10 µM each)	1.5 µl	0.3 µM each
Plasmid DNA (20 ng/µl)	1 µl	20 ng
DNA methylase (4 U/µl)	1 µl	4 U
25x SAM	2 µl	1x
AccuPrime™ Pfx (2.5 U/µl)	0.4 µl	1 U
PCR water	to 50	

b. PCR reaction

Temperature	Duration	Number of Cycles
37 °C	12-20 min (Plasmid methylation)	1
94 °C	2 min	20
94 °C	20 sec	
57 °C	30 sec	
68 °C	30 sec/kb of plasmid	
68 °C	5 min	1
4 °C	hold	1

c. Recombination reaction

This step allows the conversion of the plasmid from a linear to a circular form.

In a 1.5 ml tube add: 4 µl of reaction buffer (5x), 2 µl of enzyme mix (10x), 4 µl of PCR sample and complete to 20 µl with RNase-free water. Incubate 10 min at room temperature (RT). Stop the reaction with 1 µl of 0.5 M EDTA. The next step should be started immediately after recombination. Prepare for the transformation step before beginning the recombination reaction.

d. Transformation

- i. Thaw vials of DH5α cells on ice for 5 to 7 min.
- ii. Transfer 2 µl of the recombination reaction directly to the vial of DH5α cells and mix by tapping gently (Reaction A).
- iii. Incubate the tubes on ice for 12 min.
- iv. Immediately incubate the vials at 42 °C for 30 sec (do not exceed 30 sec) and then re-incubate on ice for 2 min.
- v. Add 250 µl of SOC medium to each vial (Reaction A) and then incubate at 37 °C for 1 h with 225 rpm shaking (Reaction B).
- vi. Add 10 µl from Reaction B to 90 µl of SOC medium, then spread 100 µl onto LB plates and incubate overnight.
- vii. Select 5 to 10 colonies and isolate DNA from the bacteria.
- viii. Analyze and confirm, by DNA sequencing, the presence of the induced mutations in the promoter.
- ix. Store DNA at -20 °C.

3. Transient transfection

- a. Culture HEK293T cells in DMEM media supplemented with 10% FBS (without antibiotics).
- b. Use trypsin to dissociate HEK293T cells (1 ml per 10 cm dish) and incubate at 37 °C for 2 min. Add DMEM and count live cells (using trypan blue) in suspension and then plate 50,000 cells per well in 12-well plates. Add a total of 1 ml of DMEM 10% FBS media to each well.
- c. Begin the transient transfection the following day.
- d. For transfection, prepare plasmids in 1.5 ml tubes as for the following example:

**Table 1. Plasmid preparation for transient transfection into HEK293T cells**

Plasmids	Condition			
	Control	MYC	MAD	MYC-MAD
<i>ITGA1</i> wild type	100	100	100	100
MYC	X	100	X	X
MAD	X	X	100	100
Empty vector	200	100	100	X
pGL4.13	2	2	2	2

**Table 2. Plasmid mix for the comparison between wild type and mutated *ITGA1* promoter**

Plasmids	<i>ITGA1</i> wild type		<i>ITGA1</i> mutated	
	Empty Vector	MYC	Empty Vector	MYC
<i>ITGA1</i> wild type	100	100	X	X
<i>ITGA1</i> mutated	X	X	100	100
MYC	X	100	X	100
Empty vector	100	X	100	X
pGL4.13	2	2	2	2

Plasmid quantities are in ng

X: no plasmid transfection

Transfect with the same total quantity of plasmid for every well.

Amounts are for one well, calculate quantities for triplicates.

- e. Add 50  $\mu$ l of buffer EC and 1.5  $\mu$ l of enhancer, mix gently and incubate for exactly 5 min.
- f. Add Effectene reagent (2  $\mu$ l/well), mix gently and then incubate for 10 min.
- g. Change the media in the 12-well plates and then add the plasmid mix to each well. Incubate for 48 h.
- h. Add one volume of passive lysis buffer to four volumes of distilled water.
- i. Remove the media from the wells and rinse once gently with cold 1x PBS.
- j. Add 250  $\mu$ l of 1x passive lysis buffer for each well of the 12-well plates and then shake gently for 15 min.
- k. Luminometer settings: set injectors 1 and 2 to dispense 50  $\mu$ l of LARII and Stop & Glo solutions.
- l. For measurement, use 10 sec for reading time and 2 sec as a delay.
- m. Transfer 10  $\mu$ l from each PLB lysate and put in 96 well-plate (avoid pipetting the cell debris).
- n. Run the measurement of luciferase activity.

### **Data analysis**

1. Response elements identified using MatInspector<sup>®</sup> software should be analyzed and selected according to their similarity to the consensus sequence. Response elements with the same sequence as the consensus response element (for the transcription factor studied) as

incomplete sequences should be taken in account for the mutagenesis analysis (Boudjadi *et al.*, 2016; Chen *et al.*, 2011; Teye *et al.*, 2008).

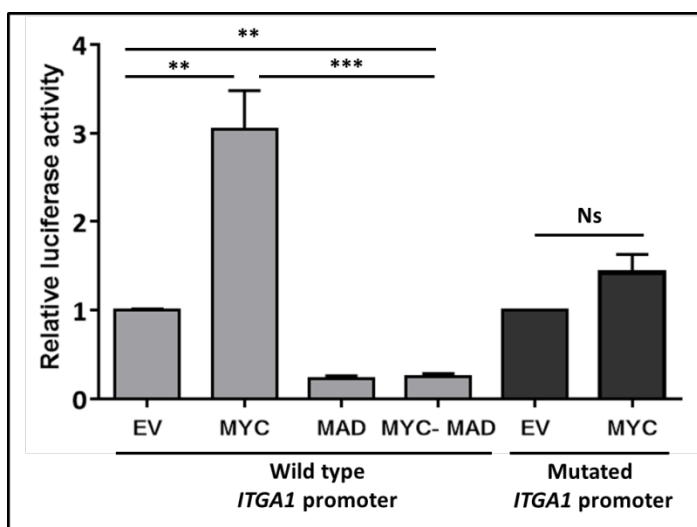
2. After plasmid mutation, the sequence of the new plasmid should be aligned to the original sequence to confirm that the introduced mutations are the only differences between the wild type promoter sequence and the mutated sequence (Figure 1).



**Figure 1. Alignment of the promoter sequences.** Comparison of the mutated *ITGA1* promoter sequence to the wild type promoter. When present, the vertical bars indicate identical nucleotides between the two sequences. The sequence of the forward primer is underlined in blue. The red square points out the mutated response element (mutation of the first response element for MYC), as reported in Boudjadi *et al.*, 2016.

3. After transient transfection and cell lysis, the enzymatic activity of luciferase (which reflects the promoter activity) will be quantified using a luminometer. The data obtained will be used to analyze the promoter activity under each condition. Two values are obtained: the first one indicates the activity of the promoter of interest and the second indicates the activity of the control promoter (reflecting the transfection efficiency). For each well, the final value is the ratio between the first and the second values.

Data may be presented as fold change as indicated in Figure 2.



**Figure2. Analysis of *ITGA1* promoter activity.** Results of the promoter gene reporter assay showing the relative luciferase activity presented as fold change compared to the empty vector

(EV). The wild type *ITGA1* promoter vector was transfected into HEK293T cells, together with the empty vector, a MYC expressing vector (MYC), the dominant negative MAD (MAD) or both (MYC-MAD). The dark gray columns indicate the activity of the mutated *ITGA1* promoter (mutation of the first response element for MYC) after transfection with the empty vector or the MYC expressing vector. Experiments were repeated three times and performed in triplicate. *t*-test. \*\* $P < 0.01$ , \*\*\* $P < 0.001$ . Ns: non-significant.

## **Notes**

1. The methylation of the template plasmid is performed in order to allow its digestion by endonucleases in the host cells after transformation. Therefore, the template will be eliminated and only the mutated plasmid will be present in the transformed cells.
2. HEK293T cells should not be allowed to reach confluency during the transfection.
3. HEK293T cells form a fragile monolayer, therefore the media change should be done in a gentle manner.
4. The MAD/MAX heterodimer recognizes the same response elements as MYC/MAX. The fact that the transfection of MAD alone or together with MYC reduced the activity of the *ITGA1* promoter is an indication that this activity was enhanced when MYC was introduced alone.
5. Store DMEM at 4 °C; heat to 37 °C before use, avoid keeping warm for a prolonged time.
6. Do not mix the bacterial vials by pipetting up and down.
7. Invert LB plate during incubation to avoid condensation and contamination.
8. Prepare plasmids at 100 ng/μl to facilitate the transfection procedure.

## **Recipes**

1. DMEM culture medium for HEK293T cells  
DMEM  
10% FBS  
2 mM GlutaMAX  
10 mM HEPES
2. LAR II solution  
Resuspend lyophilized Luciferase Assay substrate in Luciferase Assay buffer II (105 ml)  
Aliquot and store at -20 °C
3. 1x Stop & Glo  
Dilute the appropriate amount of Stop & Glo substrate (stock 50x) in Stop & Glo reagent (Always prepare a fresh solution)

## **Acknowledgments**

We thank Elizabeth Herring for reviewing the manuscript. This work was supported by the Canadian Institute of Health Research Grant MOP-123415 (JFB is member of the Centre de Recherche of the Centre Hospitalier Universitaire de Sherbrooke funded by the Fonds de la Recherche du Québec-Santé. This protocol was adapted from Boudjadi *et al.* (2016), originally published in *Oncogene*.

## **References**

1. Boudjadi, S., Carrier, J. C., Groulx, J. F. and Beaulieu, J. F. (2016). [Integrin  \$\alpha1\beta1\$  expression is controlled by c-MYC in colorectal cancer cells.](#) *Oncogene* 35(13): 1671-1678.
2. Chen, Y., Xu, J., Borowicz, S., Collins, C., Huo, D. and Olopade, O. I. (2011). [c-Myc activates BRCA1 gene expression through distal promoter elements in breast cancer cells.](#) *BMC Cancer* 11: 246.
3. Ni, H., Dydensborg, A. B., Herring, F. E., Basora, N., Gagne, D., Vachon, P. H. and Beaulieu, J. F. (2005). [Upregulation of a functional form of the  \$\beta4\$  integrin subunit in colorectal cancers correlates with c-Myc expression.](#) *Oncogene* 24(45): 6820-6829.
4. Teye, K., Okamoto, K., Tanaka, Y., Umata, T., Ohnuma, M., Moroi, M., Kimura, H. and Tsuneoka, M. (2008). [Expression of the TAF4b gene is induced by MYC through a non-canonical, but not canonical, E-box which contributes to its specific response to MYC.](#) *Int J Oncol* 33(6): 1271-1280.