Optimization of CEST reporter genes with a genetic programming Protein Optimization Evolving Tool

Or Perlman[1], Alexander R. Bricco[2], E. Alejandro Castellanos[2], Iliya Miralavy[3], Shaowei Bo[4], Tyler Gallagher[5], Leo L. Cheng[1], Michael T. McMahon[4], Wolfgang Banzhaf[3], Hiroshi Nakashima[5], Assaf A. Gilad[6,7], and Christian T. Farrar[1]

[1]Martinos Center for Biomedical Imaging, Massachusetts General Hospital and Harvard Medical School, Charlestown, MA; [2]Department of Biomedical Engineering, Michigan State University, East Lansing, MI; [3]Department of Computer Science & Engineering, Michigan State University, East Lansing, MI; [5]Department of Neurosurgery, Brigham & Women's Hospital, Boston, MA; [4]Department of Radiology and Radiological Sciences, Johns Hopkins University School of Medicine, Baltimore, MD; [6]Department of Chemical Engineering and Materials Science, Michigan State University, East Lansing, MI; [7]Department of Radiology, Michigan State University, East Lansing, MI.

Introduction: Cell and viral based therapeutics hold great promise for revolutionizing the treatment of many diseases. However, the optimization of such biological therapies and assessment of their efficacy depends critically on the ability to monitor the spread and persistence of the therapeutic agent. We have previously demonstrated that a Lysine-Rich Protein (LRP) Chemical Exchange Saturation Transfer (CEST) MRI-based reporter gene[1] can be used for imaging oncolytic virotherapy[2] as well as cardiac gene transfer therapy[3]. However, the detection sensitivity was limited, and the reporter gene was not stable due to the highly repetitive nature of the DNA sequence. To optimize the design of CEST reporter proteins we have developed a novel genetic programming algorithm, Protein Optimization Evolving Tool (POET)[4]. Here we characterize CEST peptides optimized using POET and use the best performing peptides to construct new CEST reporter genes.

Methods: POET was used to optimize the design of 12mer CEST peptides to provide maximal CEST contrast at either 3.6 or 5 ppm[4]. Each generation of optimized peptides were synthesized, and the proton Magnetization Transfer Ratio Asymmetry (MTR_{asym}) was experimentally measured on a 7T Bruker MRI scanner. The experimental results were then used to improve the predictive model of the POET. The amide exchange rates and proton volume fractions were quantified using QUESP[5] on a 14T Bruker NMR spectrometer. The contribution of the different exchangeable proton pools (amide, amine, and hydroxyl) to the overall CEST Z-spectrum was quantified by 4-pool Lorentzian fitting[6] of the Z-spectra acquired at 14T. Different combinations of the 5 best peptides were then used to generate new optimized reporter genes. Protein structures were predicted using Robetta and the 2 most disordered proteins (ASM1 and ACF2) were selected. Genes encoding these proteins and a P2A tag for western blot detection were engineered into a pcDNA3.1 P2A-eGFP vector and transfected into HEK293 cells. Western blots of HEK293 cell lysates were run to characterize the new reporters.

Results and Discussion: After 10 generations we have obtained a group of peptides with amide CEST contrast up to 4 times greater than poly-L-lysine (K12), the basis for the original LRP reporter protein. Most of the POET optimized peptides showed improvement in the average amide proton exchange rate and amide CEST contrast with some of the peptides demonstrating exchange rates almost 2-times faster than K12 (***Table 1***). Some of the POET optimized peptides had lower amide exchange rates relative to K12, however, the amine CEST contrast at 2 ppm was significantly greater for these peptides indicating that the increased MTR_{asym} at 3.6 ppm for these optimized peptides also has contributions from the amine exchangeable protons. Thus, POET can optimize and exploit both amide and amine protons to maximize the MTR_{asym} at 3.6 ppm. The reporter genes AMS1 and ACF2, consisting of different combinations of the 5 best CEST peptides, were chosen due their predicted disordered structure (***Figure 1, left***), which provides optimal solvent access for the exchangeable protons. Western blots of HEK293 cell lysates (***Figure 1, right***) show a single well-defined band at approximately the correct molecular weight indicating expression of full-length protein and a stable reporter gene. Experiments are currently underway to test the *in vitro* and *in vivo* performance of the optimized reporter genes in tumor cell lines engineered to express the new reporters.

Table 1:	Normalized CEST Contrast (%)[1]			k_{ex} (Hz)
Peptide Sequence	**Amide**	**Amine**	**OH**	**Amide**
KKKKKKKKKKKK	2.18	0.04	0.0	423
KPWHGCASRTKR	4.28	4.95	6.41	548
DKVCKIQKRKWH	2.91	1.74	0.0	422
KKRLHWIRWHCG	2.27	4.67	0.0	195
CCWHNPKWRRTR	2.05	5.23	7.19	433
KYTKTRKQSSKA	5.42	2.19	4.71	754
NSSNHSNNMPCQ	5.26	0.91	1.40	822

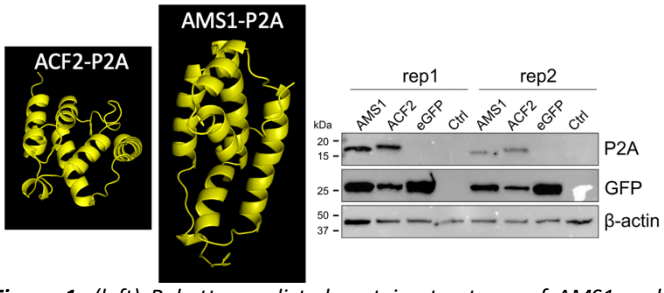

Figure 1: *(left) Robetta predicted protein structures of AMS1 and ACF2 reporter proteins; (right) Western blot of HEK293 cell lysates transfected with the reporter genes.*

References: 1. Gilad, A. A. *et al. Nat Biotech* 25, 217–219 (2007); **2**. Farrar, C. T. *et al. Radiology* 275, 746–754 (2015); **3**. Meier, S. *et al. Sci Rep* 8, 4638 (2018); **4**. Bricco, A. R. *et al. bioArXiv* (2022) doi:10.1101/2022.03.05.483103; **5**. McMahon, M. T. *et al. Magn Reson Med* 55, 836–847 (2006); **6**. Zaiss, M., Schmitt, B. & Bachert, P. *J Magn Reson* 211, 149–155 (2011).