Design of a biological half adder

M. Terzer, M. Jovanovic, A. Choutko, O. Nikolayeva, A. Korn, D. Brockhoff, F. Zürcher, M. Friedmann, R. Schütz, E. Zitzler, J. Stelling and S. Panke

Abstract: The building of complex systems from basic logic gates is one of the hallmarks of circuit design in electrical engineering. The question arises whether a similar strategy can be adopted for the design of artificial biological systems. In this paper, we present the design of two logic gates, a biological AND and a biological XOR. They can be combined to produce a half-adder, one of the fundamental elements of complex systems engineering, and represent a promising basis for the design of more complex genetic circuits. Design space exploration allowed us to screen gate variants, while sensitivity analysis of refined models contributed to the specific implementation of the gates at the DNA level. The XOR gate is based on two specific proteases, which reciprocally inactivate co-synthesised transcription factors. The AND gate is designed such that, in the presence of two signals, a tRNA suppresses the premature termination of T7 RNA polymerase translation. Computer models confirmed that both designs allow gate behaviour that is reasonably close to idealised gates.

1 Introduction

Naturally occurring biological systems integrate complex sets of information and process these to perform complicated actions. In contrast, current artificial biosystems such as oscillators, switches, or logic gates [1] are rather simple and have limited functionality. However, simple functions such as logic gates can be used as modular building blocks. This approach has allowed engineers to produce digital electronic systems with the enormous complexity of modern microprocessors. Consequently, we aimed at designing and implementing a more complex biological functionality from simple logic gates. One standard and versatile function that is well established in electric circuit engineering is the half-adder that is itself composed of two logic gates, an AND and an XOR gate. Such a half-adder leads to one output if one of two signals is present and to another output if both signals are present simultaneously. In this communication, we report the design of an artificial biological counterpart.

2 System overview

An electronic half-adder accepts two single digit binary inputs and produces a two bit-sum as the output. The more significant bit of the output is called the carry-over, in analogy to the addition procedure on paper, where this number is carried over to the next position. In the binary half-adder, three cases can be distinguished: first, both inputs can be 0. In this case, the sum is 00 and the carry-over is 0. In the second case, one of the inputs is 1 and the other is 0, so the sum is 01 and the carry-over is still 0. Finally, if both inputs are 1, the sum is 10, and the carry-over is 1.

doi:10.1049/iet-stb:20070013

Paper first received 19th February 2007

M. Terzer, A. Choutko, O. Nikolayeva, A. Korn, D. Brockhoff, F. Zürcher, M. Friedmann, R. Schütz, E. Zitzler, J. Stelling, and S. Panke are with ETH Zurich, CH-8092 Zurich

M. Jovanovic is with University of Zurich, CH-8057 Zurich

M. Terzer, M. Jovanovic, A. Choutko, O. Nikolayeva, A. Korn and D. Brockhoff contributed equally to the work

E-mail: panke@ipe.mavt.ethz.ch

The less significant bit of the sum is the result of an XOR logical operation on the inputs, resulting in 1 only when exactly one of the operands is 1 and 0 otherwise. The more significant bit of the sum is the outcome of an AND operation on the inputs, which is 1 only when both inputs are 1 and 0 otherwise. The constituent gates of the half-adder are operating in parallel, thus the implementations can be designed independently.

In a biological half adder, any two stimuli representing the inputs, for example, presence of light and/or chemicals, are converted into two biological signals (e.g. enzyme activity or regulatory gene expression), which then are processed by the system. The outcome of this is either the absence of a reporter protein, RFP synthesis (one signal present, value of 1 in the output of the XOR gate), or GFP synthesis (both signals present, value of 1 in the AND output) (Fig. 1*a*). Depending on the chosen implementation and integration of the gates, additional measures might be necessary to connect a single input to two logic gates. The signal duplexer serves this purpose in our design.

3 Exploration of design space

The advantages of mathematical modelling include the identification of parts that contribute most to properties of interest and faster and cheaper generation and testing of hypotheses [2]. Successful system design requires several iterations of modelling, testing, model selection, and refinement. Early iterations focus on design space exploration (Fig. 1*b*), where the number of system design alternatives is reduced based on several cycles of simulation and optimisation—and of course the consideration of biological feasibility. At later stages, models can be further improved by accounting for experimental data. We focus here on the design space exploration.

The modelling was carried out separately for the XOR and the AND part. Various concepts to design biological XOR or AND gates are theoretically possible, a selection of which is collected in Fig. 2. Simulations of these models were performed to identify a suitable steady-state behaviour. Furthermore, the models were analysed for parameter sensitivity to identify relatively robust designs.

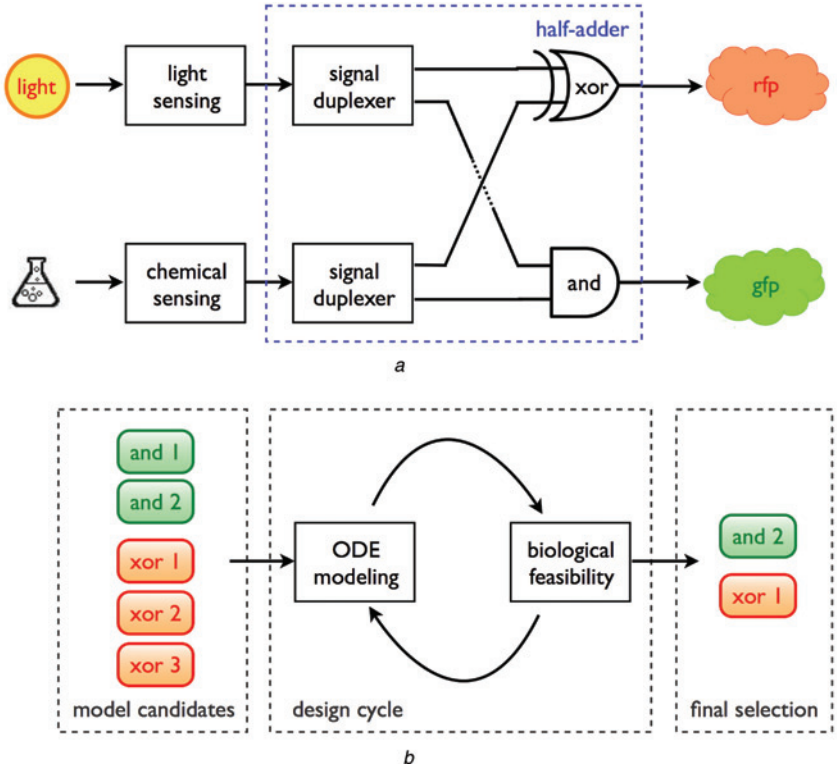

Fig. 1 *Schematic overview of a biological half-adder system*

a Two biological stimuli, light and a chemical signal, are sensed by respective devices. The duplexers direct the input signals to XOR and AND gates. In the case when one of the two inputs is present, the XOR gate is activated and produces RFP. When both inputs are present, the AND gate is activated and produces GFP. Lack of both inputs results into the absence of both reporter proteins

b Engineering design approach to model selection: several AND and XOR model candidates are generated, based on abstract notions of possible implementation. Through iterations of ODE modelling and biological feasibility considerations, relatively better models are selected to be experimentally tested

Finally, the models were analysed for feasibility of biological implementation. The corresponding MATLAB scripts are available from the authors upon request.

3.1 ODE systems

Ordinary differential equations (ODEs) were used for modelling the gate variants. Altogether, the models contain four different types of ODEs describing enzymatic transformation, constitutive transcription, regulated transcription, and translation (see online supplementary material). For example, the equations below illustrate the system of ODEs for the final AND gate

$$\frac{d[T7Pol]}{dt} = k(tl)_{T7Pol}[mT7Pol] \times [tRNA] - d_{T7Pol}[T7Pol]$$

$$\frac{d[T7Pol^*]}{dt} = k(tl)_{T7Pol^*}[mT7Pol] - d_{-T7Pol^*}[T7Pol^*]$$

$$\frac{d[tRNA]}{dt} = k_{tRNA}[Input1] - d_{tRNA}[tRNA]$$

$$\frac{d[mT7Pol]}{dt} = k_{mT7Pol}[Input2] - d_{mT7Pol}[mT7Pol]$$

$$\frac{d[mOut]}{dt} = p_{Out} + \frac{k(tr)_{Out}\left(\frac{[T7Pol]}{K_{Out}}\right)^{n_{Out}}}{1 + \left(\frac{[T7Pol]}{K_{Out}}\right)^{n_{Out}}}$$

$$\frac{d[Out]}{dt} = k(tl)[mOut] - d_{Out}[Out]$$

The bracketed species above refer to concentrations, and the symbols are detailed in the legend of Fig. 3.

3.2 System selection

Fig. 2 shows surface plots for the different models corresponding to different design alternatives. The output rates for the gates as a function of two inputs at steady state after 12 h simulation are shown. The left and the middle column represent models that were prepared early in the model refinement process. These guided decision making and have little more than qualitative value. The right column shows final versions. These are more realistic since experimental data, available from other systems, has been used to estimate parameters such as reaction constants [3, 4].

Fig. 2*a*–*c* shows three versions of the XOR gate. An ideal XOR gate should have high output if one input is high and low output if both inputs are either low or high. Output behaviour of XOR 1 (Fig. 2*a*) is clearly preferable to the behaviour of XOR 2 (Fig. 2*b*). Combined with a simpler biological implementation, this led us to select XOR 1 for elaboration. XOR 3 (Fig. 2*c*) is based on XOR 1 and has been adapted to the selected biological realization (see below) and refined with realistic parameter values.

Fig. 2*d*–*f* show AND gate versions. An ideal AND gate has high output only if both inputs are high. AND 1 (Fig. 2*d*) shows AND-like behaviour, but the thresholds for low and high output are very close and might be impossible to distinguish in experiments. On the other

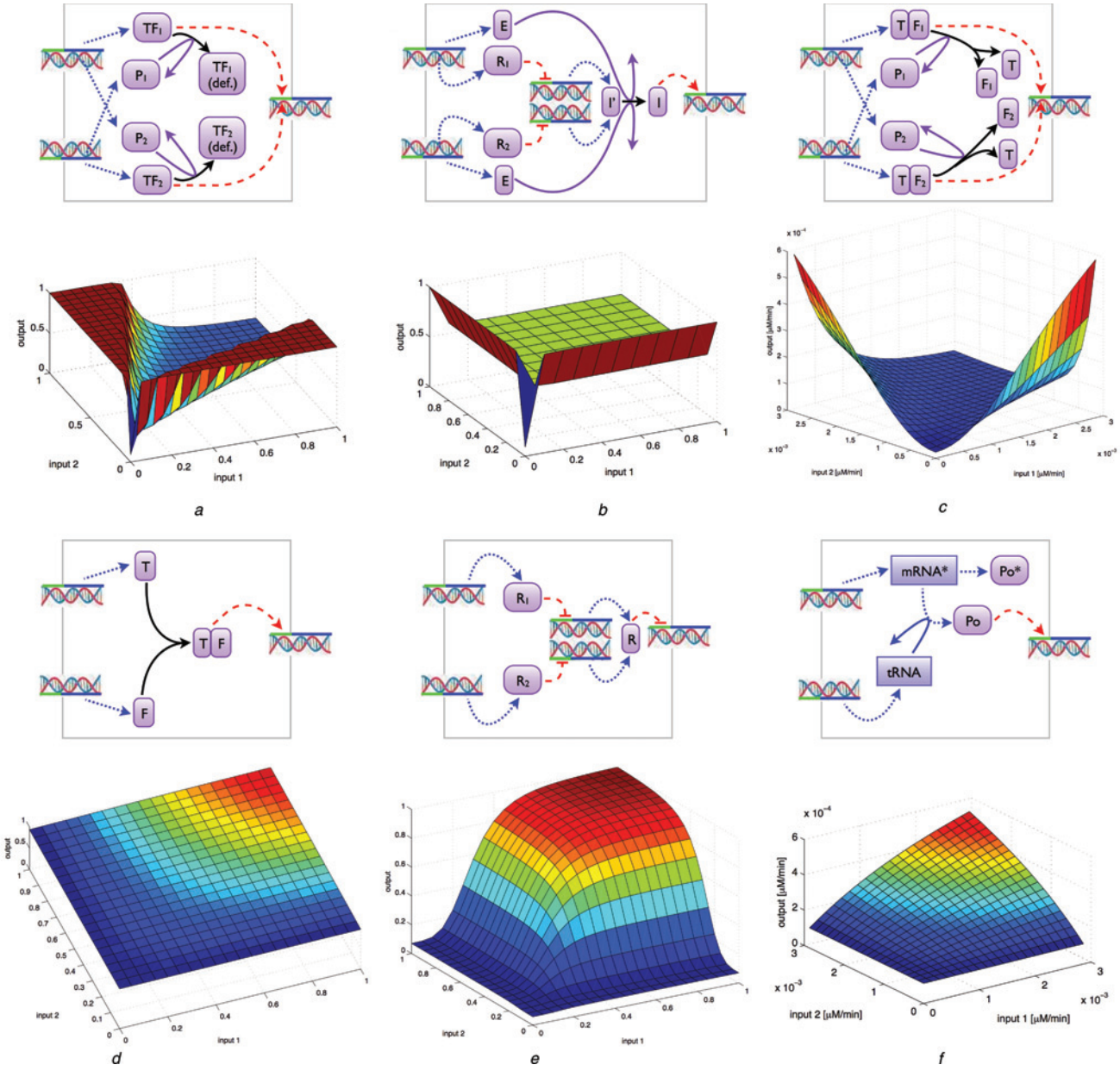

Fig. 2 *Biological structure and modelling results for various design alternatives*

The modelling results in the second and forth row illustrate the steady-state output of the system (z-axis) as a function of two varying inputs (x- and y-axis) for the biological system illustrated above. The green parts of the DNA elements represent regulatory, the blue part coding sequences. Dotted blue arrows indicate protein expression; solid black arrows represent enzymatic reactions, enzyme participation shown in purple. Dashed red lines highlight regulatory interaction (arrow for induction, bar-headed for repression). Early stage models (*a*, *b* and *d*, *e*) have mainly been used to support decision making and thus have little more than qualitative grade. For this reason, normalised values are shown on the axes

a XOR 1: Input 1 triggers the synthesis of transcription factor (TF) 1 and protease (P) 2 from the upper DNA element. TF1 triggers synthesis of the reporter (XFP). Input 2 triggers synthesis of TF2 and P1 from the lower DNA element and also leads to synthesis of XFP. When both inputs are present simultaneously, P1 digests TF1 and P2 digests TF2, resulting in mutual inactivation of the transcription factors

b XOR 2: Either input triggers synthesis of a repressing TF R1 or R2 and of an enzyme E. E is needed to activate an inducing TF I' to I, and I triggers the synthesis of XFP. Only the presence of both repressors, R1 and R2, stops the synthesis of I' and prevents synthesis of XFP

c XOR 3: Final XOR model derived from XOR 1 using two very similar transcription factors, TF1 and TF2, with corresponding specific proteases P1 and P2, which inactivate the transcription factors by mutual cleavage

d AND 1: Two parts T and F of a transcription factor triggering synthesis of XFP are produced from two DNA elements. Both input signals are required for a functional TF

e AND 2: (double repression): XFP synthesis is repressed in the presence of repressor R. R can be synthesized from two DNA elements, and repression of R synthesis requires another repressor for each element, R1 and R2. Consequently, only the presence of two inputs turns on XFP synthesis

f AND 3: Input 1 triggers the synthesis of a truncated polymerase Po* from mRNA*. Input 2 triggers the synthesis of a suppressor tRNA that allows to suppress the early termination of translation. Simultaneously, inputs 1 and 2 allow the production of a full-length polymerase to trigger synthesis of XFP from its cognate promoter

hand, AND 2 (Fig. 2*e*) has almost ideal characteristics but is complex to implement biologically. AND 3 has high output for high inputs, but the transition from low to high output is not as steep as for AND 2. Finally, the possibility of simple and elegant implementation lead us to select AND 3.

4 Biological implementation

4.1 XOR gate

As already mentioned above, an XOR gate produces an output only if exactly one of the two possible input signals is present. This requires a system where gene

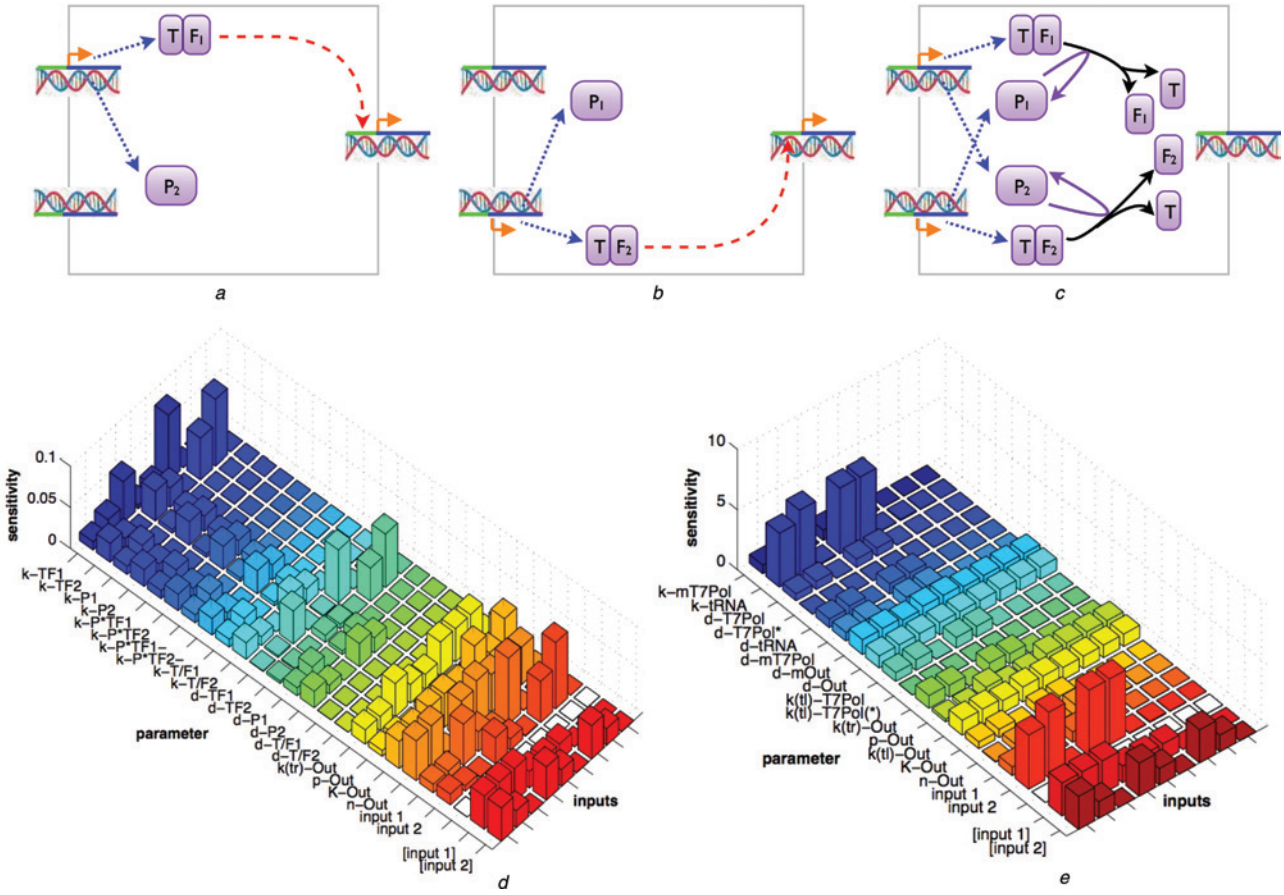

Fig. 3 *XOR gate under different signal conditions (a–c); sensitivity plots of the final XOR and AND gates (d–e)*

Notation as in Fig. 2; orange arrows represent triggered expression

a only Signal 1 is present—due to the incompatibility of the expressed transcription factor and the protease, reporter gene expression is induced

b only Signal 2 is present—same outcome as in *a*

c Signal 1 and Signal 2 are present—due to mutual cleavage of the transcription factors by the corresponding proteases, reporter gene expression is not induced

d–e The z-axis shows the normalised sensitivities. Higher bars represent stronger output sensitivity to changes of corresponding parameters. The other axes correspond to the parameters (longer axis) and nine different input combinations (all combinations of low, medium and high activity for both inputs). The two right most series [(input 1) and (input 2)] show which specific combination of inputs has been used. For example, the first row describes high input 1 and high input 2, the next high input 1 and middle input 2, etc. Note that some directly output-related parameters of the XOR gate are not displayed to enhance clarity for other parameters more relevant for the present work

The following abbreviations have been used:

(i) prefixes identify the parameter type: d = degradation constant; k = reaction rate constant; k(tl) = kinetic constant of translation; k(tr) = kinetic constant of transcription; K = Hill constant; n = Hill coefficient; p = constant background level (leakiness) of regulated transcription, typical values are 5–10 % percent of k(tr)

(ii) suffixes identify following terms: mOut = mRNA encoding the output signal (reporter); mT7Pol = mRNA encoding the T7 polymerase; Out = output signal (reporter protein); P1 and P2 = Protease; T7Pol = full length T7 polymerase; T7Pol* = truncated T7 polymerase; TF1 and TF2 = transcription factor; T/F1 and T/F2 = cleaved transcription factor; P*TF1 and P*TF2 = enzyme substrate complex in the forward direction; P*TF1- and P*TF2- = enzyme substrate complex in the reverse direction; tRNA = suppressor tRNA

expression triggered by signal 1 inhibits the gene expression triggered by signal 2 and vice-versa. We designed a system consisting of two biological units: (i) a bi-modular transcription factor with the DNA binding domain separated from the transcriptional activation domain by a protease cleavage site; and (ii) the corresponding protease with a high specificity for the protease cleavage site. Presence of the protease should lead to efficient cleavage and therefore deactivation of the transcription factor. Therefore, a functional XOR gate requires two transcription factors acting on the same promoter, but separated by different protease cleavage sites corresponding to two different and specific proteases. Signal 1 alone leads to expression of the reporter gene (Fig. 3*a*): First, the bi-modular transcription factor 1 with cleavage site 1 is synthesised together with protease 2. However, protease 2 cannot cleave transcription factor 1, so the downstream reporter is synthesised. The same principle applies if only signal 2 is present (Fig. 3*b*). This triggers synthesis of transcription factor 2, containing the protease cleavage site 2 and protease 1. In analogy to the situation described above, transcription factor 2 is not cleaved and induces reporter gene expression. If both signals are present (Fig. 3*c*), the synthesis of both variants of the transcription factor and of both proteases is triggered, which leads to cleavage of both transcription factors and the reporter gene is not expressed.

Implementation of this system requires two highly specific proteases which can be synthesised in their active form in bacteria *in vivo* and preferably have no additional cleavage sites in other proteins of the host, such as the proteases from the tobacco etch virus (TEV) [5] and the tobacco vein mottling virus (TVMV) [3].

The bi-modular transcription factor consists of the DNA binding domain of the bacteriophage lambda repressor protein (λcI) fused to the N-terminal domain of the RNA polymerase α subunit via a linker containing the corresponding TEV or TVMV protease cleavage sites ([6] and A. Hochschild, pers. comm.). Binding of this transcription

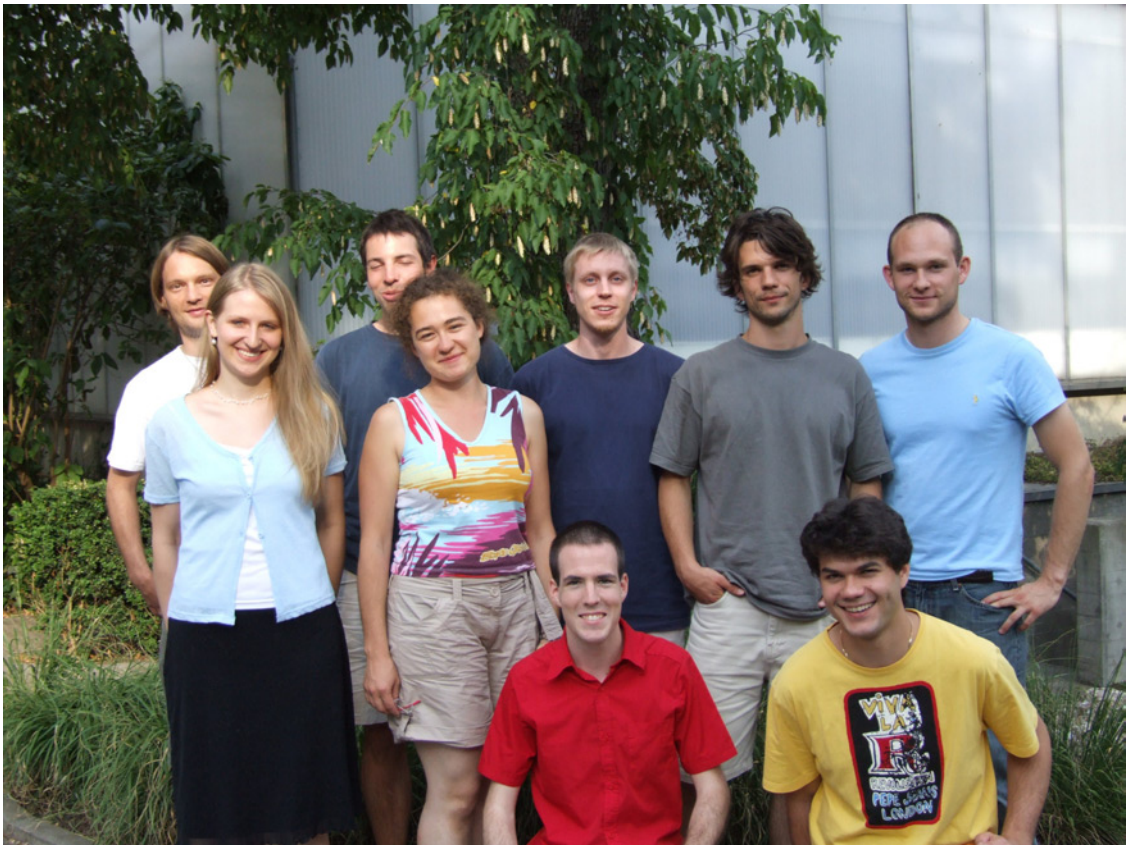

Fig. 4 *The ETH Zurich 2006 iGEM team*

factor to its cognate λ operator promotes transcription from a modified P_{lacZ} promoter and finally the synthesis of RFP as a reporter.

4.2 AND gate

An AND gate produces only an output if both input signals are present. This requirement is met by coupling the reporter gene expression to the presence of two essential biological units, each induced exclusively by the presence of one signal. In our design, signal 1 leads to synthesis of the T7 RNA polymerase, which recognises specifically its cognate T7 promoter [7]. An artificial early stop codon in the coding sequence of the polymerase leads to synthesis of a truncated, non-functional protein. Signal 2 triggers the synthesis of a suppressor tRNA that recognises the stop codon within the T7 RNA polymerase and prevents the premature termination of translation by incorporating a glutamine. Consequently, only the presence of both signals leads to a functional T7 RNA polymerase and expression of the downstream reporter gene (GFP) from the T7 promoter (Fig. 2*f*).

5 Sensitivity analysis of final designs

In silico sensitivity analysis can provide preliminary information about the expected robustness of the system in advance and help in guiding system and experimental design. We determined sensitivity as the change in output for a small change in a specific parameter or input value by computing partial derivatives of the ODEs with respect to all parameters analytically, i.e. by computing the Jacobian Matrices of the system. For details, see the online supplementary material.

The sensitivity plots confirm the model structure for both gates (Fig. 3*d*, *e*). For example, the expected insensitivity of the XOR gate output towards changes in the degradation rate of the cleaved transcription factor is confirmed (parameters d-T/F1 and d-T/F2, Fig. 3*d*), and so is the insensitivity of the AND gate's output towards the degradation rate of the truncated T7 Polymerase (d-T7Pol*, Fig. 3*e*).

Next, the sensitivity analysis indicates that the degradation rates of the proteases in the XOR gate (d-P1 and d-P2, Fig. 3*d*), have a relatively small influence on the system response, which is not obvious as the proteases act catalytically and their half-life might thus have a strong influence on the performance of the system. Guided by the analysis, we refrained from shortening the protease half-life by tagging [8, 9]. Interestingly the analysis actually showed that the degradation rates of the transcription factors themselves, d-TF1 and d-TF2, are critical for the system response.

Finally, the analysis suggests that the AND gate has a strong asymmetric behaviour regarding the two input signals (Fig. 3*e*). It shows a much higher sensitivity toward the input triggering suppressor tRNA synthesis (input 2, k-tRNA) than toward the input triggering T7 polymerase synthesis. This asymmetry is not reflected in the simulations of the steady state model (Fig. 2*f*). The reason for this will be addressed in future work.

6 Implementation

Based on the results discussed above, we designed two DNA elements and deposited them in the MIT registry of standard biological parts (http://parts.mit.edu), one for the AND gate (registry part BBa_J34100) and one for the XOR gate (BBa_J34200) and had them *de novo* synthesised. For so far unclear reasons, reassembly of the XOR gate from subfragments into high copy number plasmids has remained elusive. Next steps will include the integration of the AND and XOR gates with sensing devices, such as the light sensing device [10] (BBa_I15008, BBa_I15009, BBa_I15010, BBa_R0082) and an IPTG-sensing device based on the LacI/P_{lacZ} system [11].

7 Conclusion

The combination of artificial self-sufficient building blocks allows the realisation of more complex functionalities in biological systems, such as a half-adder from XOR and AND gates. Design space exploration coupled to biological feasibility assessment allows the rational selection of superior designs and sensitivity analysis of a refined model can help identify the most crucial influences in the system to direct experimental efforts.

8 Acknowledgments

We are indebted to the EU Project SYNBIOCOM and the ETH Zurich for funding and to Victor de Lorenzo for inspiration with the AND gate.

9 References

1 Sprinzak, D., and Elowitz, M.B.: 'Reconstruction of genetic circuits', *Nature*, 2005, **438**, pp. 443–448
2 Kell, D.B., and Knowles, J.D.: 'Modeling in Cellular Biology System' in Szallasi, Z., Stelling, J., and Periwal, V. (Eds.): 'The role of modeling in systems biology' (MIT Press, 2006), pp. 3–18
3 Nallamsetty, S., Kapust, R.B., Tozser, J., Cherry, S., Tropea, J.E., Copeland, T.D., and Waugh, D.S.: 'Efficient site-specific processing of fusion proteins by tobacco vein mottling virus protease in vivo and in vitro', *Prot. Express. Purific.*, 2004, **38**, pp. 108–115
4 Conrad, E.D., and Tyson, J.J.: 'Modeling in Cellular Biology System' in Szallasi, Z., Stelling, J., and Periwal, V. (Eds.): 'Modeling molecular interaction networks with nonlinear ordinary differential equations' (MIT Press, 2006), pp. 97–124
5 Kapust, R.B., and Waugh, D.S.: 'Controlled intracellular processing of fusion proteins by TEV protease', *Prot. Express. Purific.*, 2000, **19**, pp. 312–318
6 Hu, J.C., Kornacker, M.G., and Hochschild, A.: 'Escherichia coli one- and two-hybrid systems for the analysis and identification of protein-protein interactions', *Methods*, 2000, **20**, pp. 80–94
7 Tabor, S., and Richardson, C.C.: 'A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes', *Proc. Natl. Acad. Sci. USA*, 1985, **82**, pp. 1074–1078
8 Gottesmann, S., Roche, E., Zhou, Y., and Sauer, R.T.: 'The ClpXP and ClpAP proeases degrade proteins with carboxy-terminal peptide tails added by SsrA-tagging system', *Genes Dev.*, 1998, **12**, pp. 1338–1347
9 Herman, C., Thevenet, D., Bouloc, P., Walker, G.C., and D'Ari, R.: 'Degradation of carboxy-terminal-tagged cytoplasmic proteins by the Escherichia coli protease HflB (FtsH)', *Genes Dev.*, 1998, **12**, pp. 1348–1355
10 Levskaya, A., Chevalier, A.A., Tabor, J.J., Simpson, Z.B., Lavery, L.A., Levy, M., Davidson, E.A., Scouras, A., Ellington, A.D., Marcotte, E.M., and Voigt, C.A.: 'Synthetic Biology: Engineering *Escherichia coli* to see light', *Nature*, 2005, **438**, pp. 441–442
11 Reznikoff, W.S.: 'The lactose operon-controlling elements—a complex paradigm', *Mol. Microbiol.*, 1992, **6**, pp. 2419–2422