Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supporting Information for

Biomimetic Ultra-Broadband Perfect Absorbers Optimised with Reinforcement Learning

Trevon Badloe,[a] Inki Kim,[a] and Junsuk Rho*[a,b]

a. Department of Mechanical Engineering, Pohang University of Science and Technology (POSTECH), Pohang 37673, Republic of Korea.

b. Department of Chemical Engineering, Pohang University of Science and Technology (POSTECH), Pohang 37673, Republic of Korea

* E-mail: jsrho@postech.ac.kr

S1: Results for other materials

Along with the transition metals that were chosen to be studied due to their high imaginary refractive index, we also used the same DDQN to optimise gold (Au), silver (Ag), and silicon (Si) moth-eye structures. The results are shown in Table S1 in addition to the materials in the main text.

Material	**Periodicity (nm)**	**Height (nm)**	**Radius of Curvature**	**Spacer Thickness (nm)**	**Spacer Material**	**Substrate Material**	**Average Absorption (%)**
Chromium	*310*	*620*	*18*	*30*	SiO_2	*Ti*	*97.2*
Iron	*355*	*405*	*37*	*20*	SiO_2	*Au*	*97.3*
Nickel	*505*	*665*	*47*	*100*	Al_2O_3	*Cr*	*92.9*
Titanium	*300*	*515*	*20*	*200*	Al_2O_3	*Au*	*98.6*
Tungsten	*385*	*700*	*25*	*85*	Al_2O_3	*Cr*	*97.7*
Vanadium	*380*	*700*	*24*	*200*	SiO_2	*Ni*	*92.3*
Gold	**585**	**295**	**35**	**60**	$\mathbf{Al_2O_3}$	**W**	**85.9**
Silver	**585**	**590**	**70**	**55**	$\mathbf{Al_2O_3}$	**Ti**	**65.0**
Silicon	**605**	**575**	**64**	**150**	**Si**	**Ti**	**90.2**

Table S1: Results for extra materials that did not fit the criteria in the main paper.

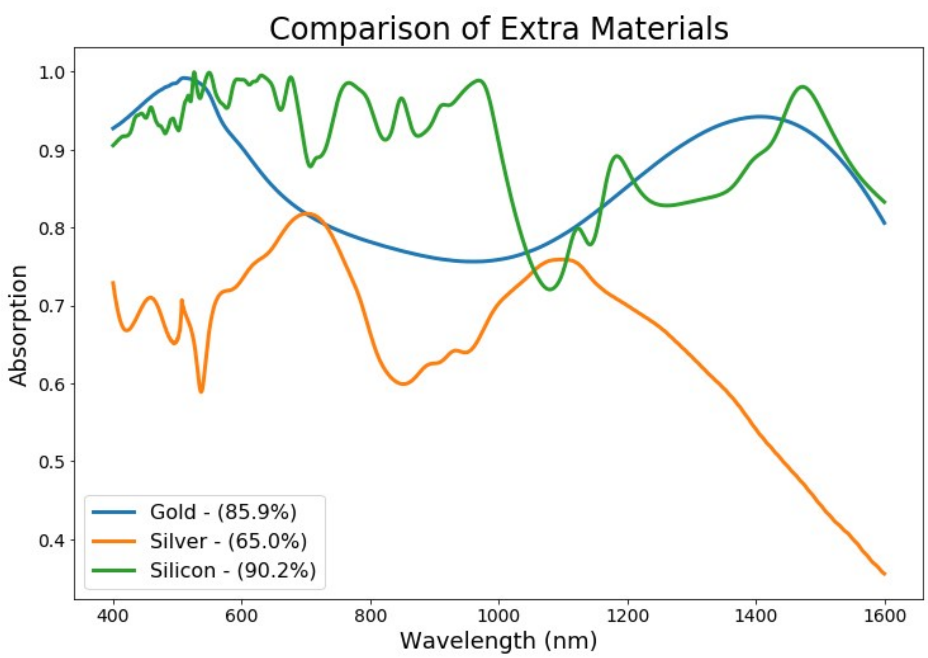

Figure S1. Absorption spectra of extra materials, gold, silver, and silicon.

Figure S1 shows the absorption spectra of the extra materials. Interestingly, the only time that the DDQN chose to use silicon as the spacer layer was in the case of the silicon moth-eye structure, creating a silicon moth-eye structure on top of a silicon layer. Since silicon is a dielectric material, the mechanism for the absorption is of course different. In Figure S2(a), we can see that the magnetic field is strongly confined inside the moth-eye structure, while at long wavelengths power is deposited in the metallic bottom layer. The gold structure is significantly different to other structures, as the radius of curvature is extremely small, creating an almost nanorod structure rather than the moth eye structure. This means that the strong field confinement in the gaps of the moth-eye structures cannot exist. In this case, the magnetic field was strongly confined under the structures in the spacer layer and that is where the power was lost, akin to a Fabry-Perot resonator as can be seen in Figure S2(b). The fields for silver are shown in Figure S2(c), giving the worst response of all the materials tested, although interestingly being the most similar to the field profiles for the materials discussed in the main text.

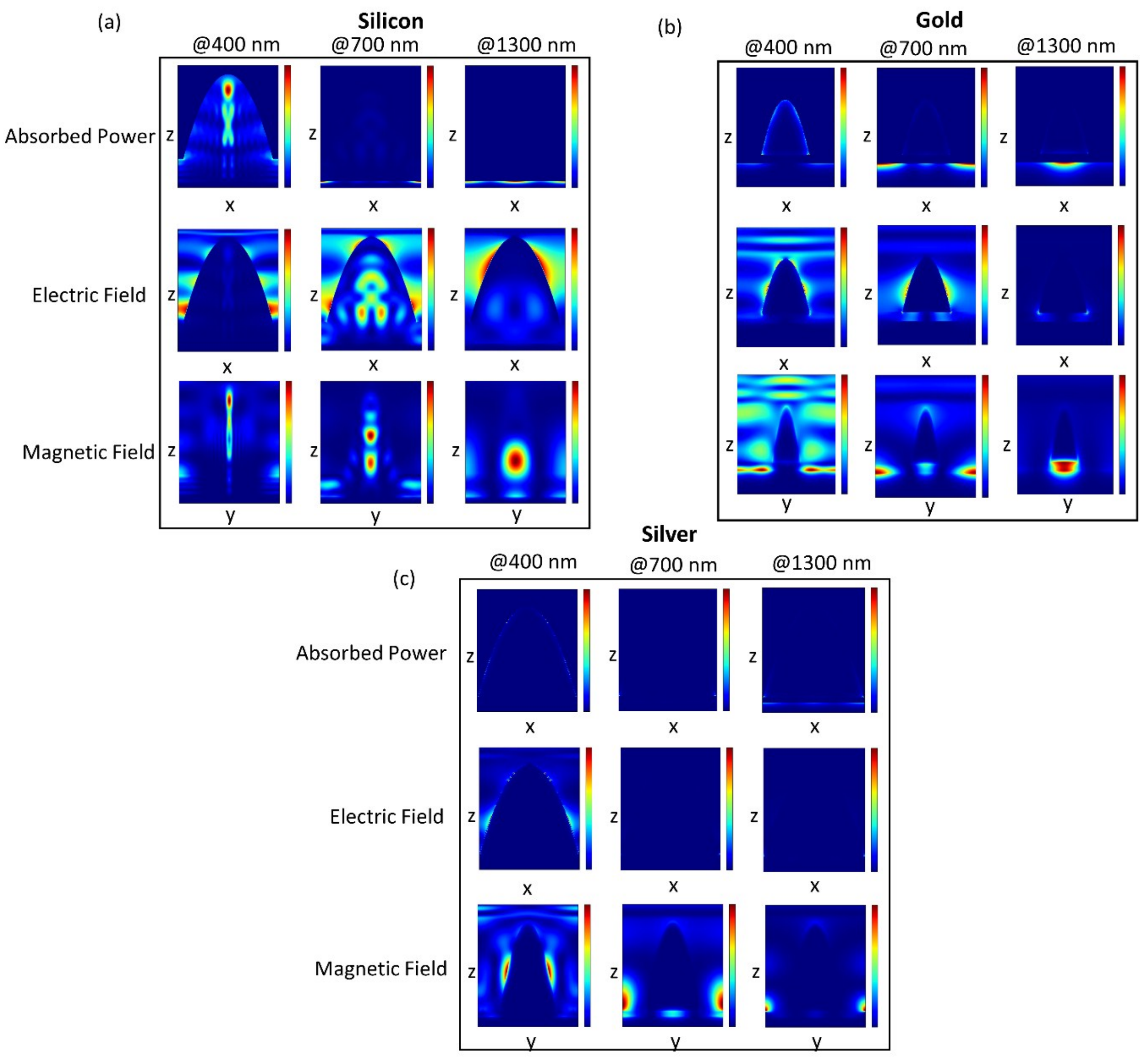

Figure S2. Field profiles for (a) gold, (b) silicon and (c), silver. The colour bars represent the minimum (blue) and maximum (red) field.

S2: Field profiles

The field profiles of the materials not shown in the main text are presented here in Figure S3.

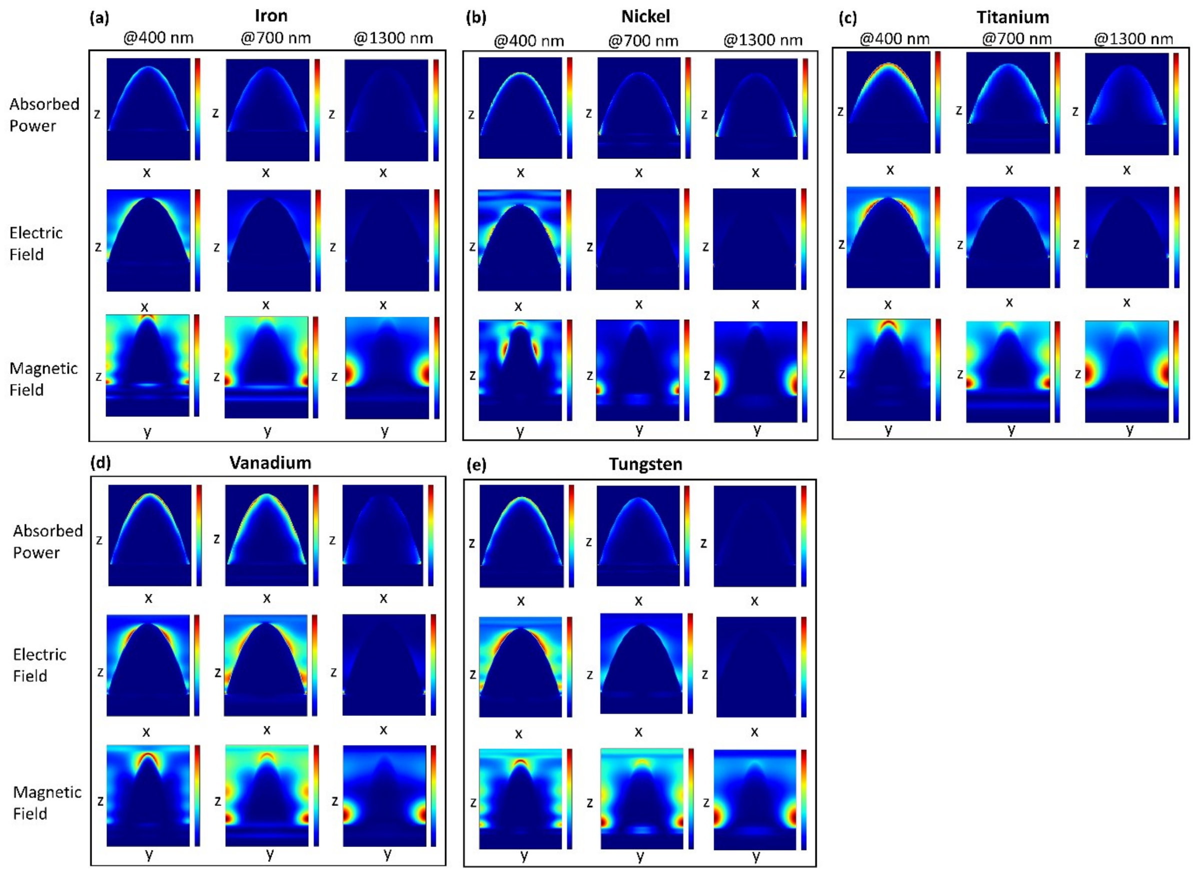

Figure S3. Field profiles for (a) iron, (b) nickel, (c) titanium, (d) vanadium, and (e) tungsten. The colour bars represent the minimum (blue) and maximum (red) field.

All of the fields have very similar features to chromium as discussed in the main paper.

S3: Structure comparison

To check whether the moth-eye structure does indeed give the best antireflection and improved absorption, we compared the moth-eye structure with similarly sized structures of different shapes for the chromium metasurface. Namely a planar surface, a nanorod, and a nanopyramid as discussed in the main paper. The results are shown in Figure S4. The moth-eye structure gives the highest overall average absorption, followed by the cone, then rod and finally the planar structure. The planar structure has low absorption since the absorption can only come from the material properties, i.e. the imaginary part of the refractive index. The cone structure has a comparable overall absorption, but still not as high as the moth-eye structure. The nanorod shape shows 3 peaks of perfect absorption, rather than broadband absorption.

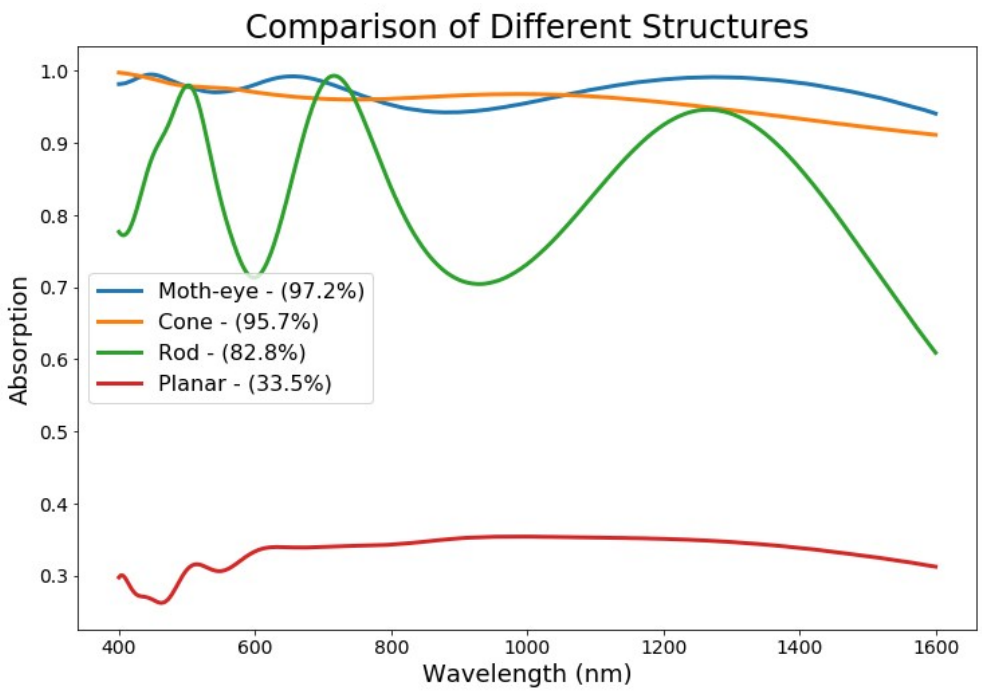

Figure S4. Comparison of the absorption from different types of structure.

S4: Epsilon greedy policy

Figure S5 shows the decaying of the epsilon greedy policy used in training the agent. A maximum of 0.95, minimum of 0.1 and a decay rate of 0.1 were chosen. For every step, a random number is generated between 0 and 1, and if it is larger than epsilon at that step, the network is used to choose the action. Otherwise a random action is taken. When epsilon is small, the network is being used up to 90% of the time, which is called exploitation. In this regime, the actions taken by the agent are decided by the target network by the action that is predicted to give the highest future reward. This means that the agent is trying to exploit the environment to get to the highest possible absorption and end the episode. When a random action is taken, this is called exploration. This way the agent can visit lots of different states at the beginning of training to gather knowledge and information on the environment to train with later. Without this stage, the agent would have no replay memory to learn from. During testing, an epsilon was set to a constant value of 0.1 to maintain that the learned policy was being used almost all the time.

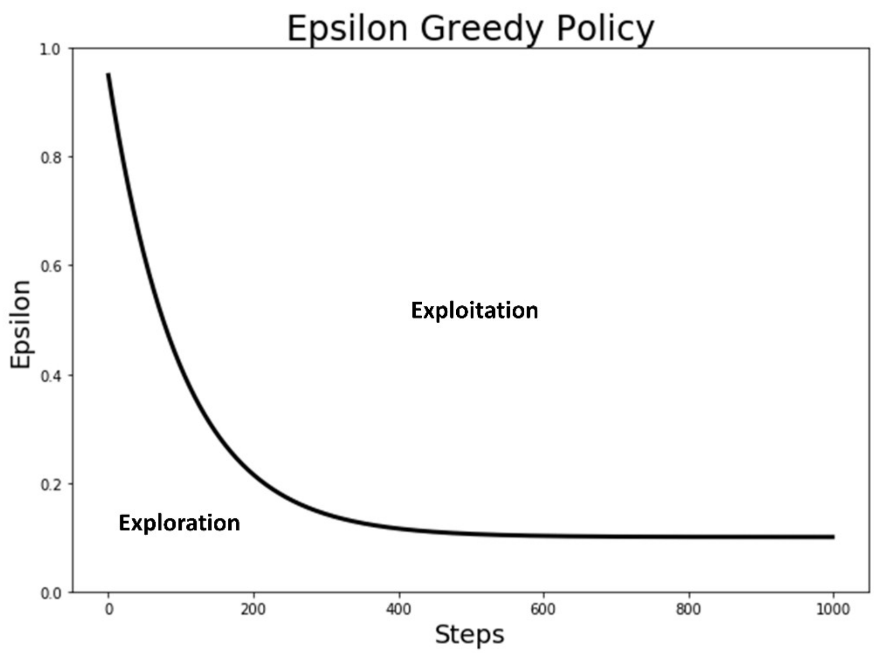

Figure S5. The epsilon greedy policy used here, with a maximum of 0.95, minimum of 0.1 and a decay rate of 0.1.

S5: Reward shaping

We were careful to design a reward system that is robust against reward hacking by the agent. Since our goal is to reach the desired absorption in the shortest possible time, we have to design the reward system in a way that encourages the agent to finish the task in as few steps as possible. A common way to achieve this is to use negative rewards. Here we shape the rewards to be maximal for absorption that is in the goal region of 90%, with a terminal state and extremely large reward. The function is shown below, with a visual representation shown in Figure S6.

$$reward = \begin{cases} -10 & if\ absorption < 85\% \\ \left(\dfrac{absorption}{90}\right)^9 - 1 & if\ absorption > 85\% \\ 10{,}000\ (and\ end\ the\ episode) & if\ absorption > 99\% \end{cases}$$

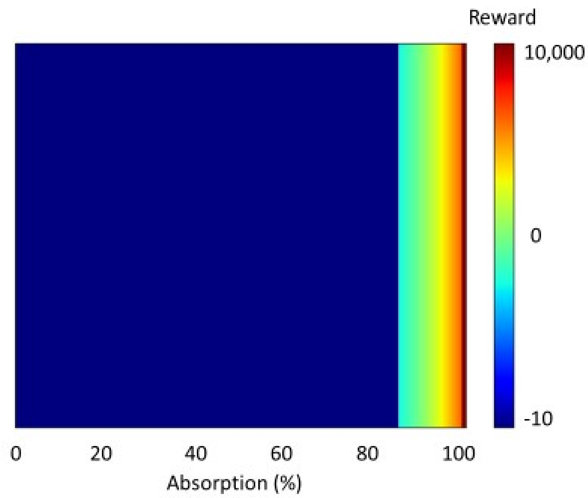

Figure S6. A visual representation of the rewards available to the agent. It receives a negative reward between 0 and -1 for absorption over 85%, and positive rewards for absorption over 90%, with a large reward of 10,000 if it reaches 99% absorption.