別紙２

# Results report

１．Title of Research and Development： Neural circuit mechanisms of reinforcement learning

２．Principal Investigator：Kenji Morita （Graduate School of Education, The University of Tokyo, Lecturer）

３．Counterpart Principal investigator：Abigail Morrison （Jülich Research Centre, Professor（Germany））

４．Results of Research and Development:

(1) Models of corticostriatal CCS and CPn cells

Estimations of the parameters of the computational models of crossed-corticostriatal (CCS) and corticopontine (CPn) cells (on the NEURON simulation environment) based on the detailed anatomical and physiological data (obtained by Dr. Morishima and Dr. Kawaguchi) have been done by Dr. Kondo, and parameter sets that could reproduce the experimental results have been obtained. Then, using the models with the obtained parameter sets (two for each cell type), simulations of the cases where the cells receive synaptic inputs that oscillate as a population with various frequencies have been done by Dr. Kondo, and conditions for the generation of "doublets" (two action potentials with short intervals) have been clarified.

(2) Model of cortico-basal ganglia circuit

Simulations of the cases where the direct or indirect pathway of the basal ganglia is blocked in the cortico-basal ganglia circuit have been elaborated, and it has been shown that the hypothesis that these two pathways represent reward predictions with time difference (Morita et al., 2012, Trends Neurosci) could account for the experimental results that the blockade of the direct or indirect pathway respectively impaired learning based on positive values (or prediction errors) and the initial phase of reversal learning, which is considered to be based on negative prediction errors. The manners the blockade of each basal ganglia pathway relates to addiction have also been examined through simulations, and it has been shown that these manners can significantly depend on the degree of exploration (over exploitation) upon choice.

(3) Model of the temporal changes of dopamine concentration

Model of the cortico-basal ganglia circuit implementing reinforcement learning with the assumption that dopamine represents reward prediction error has been elaborated by incorporating the assumption that plastic changes of the strengths of cortico-striatal synapses decay in time (by Kato and Morita). It has been shown that various temporal patterns of dopamine signals can be reproduced by the model. Also, based on simulations and analyses of the model, we have proposed a circuit/synapse-level mechanism of the ever-suggested link between dopamine and motivation (paper under review).

(4) Explorations of integrative mechanisms

Reviews of the studies that have been conducted by the Japanese and German groups, including the abovementioned studies, and also related studies done by other researchers on the neural circuit mechanisms of reinforcement learning, have been done by Morita and the member of the German group (Dr. Morrison and Dr. Jitsev) with a special focus on value-based action selection (decision making), and have been published as a joint review article. In the article, we have proposed hypotheses on how reinforcement learning algorithms (Q-learning and SARSA) can be implemented in the neural circuits, as well as on potential computational roles of the complex patterns of cortical and striatal neural activity suggested in recent experimental and theoretical studies.