Chapter 1: Introduction

Nearly all nations worldwide have been affected by the enormity of Coronavirus disease 2019 (COVID-19), including Indonesia. As of 10 September 2021, the confirmed COVID-19 cases in Indonesia have hit around 4 million cases with more than 100 thousand deaths, as stated by the World Health Organization (2021). This ongoing outbreak is generated by a virus infection, which is the severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2). The virus itself is immensely infectious and has a 5-6 days incubation period, which explains the fast spreading of the disease within a short time. Furthermore, SARS-CoV-2 also has several transmission routes, including droplets, contacts, aerosol, urine, fomites, and mother-to-child transmission (Yang et al., 2020; Zheng, 2020).

SARS-CoV-2 is part of the *Coronaviridae* family and is classified as a positive-sense single-stranded enveloped RNA virus. The genome architecture of SARS-CoV-2 consists of 14 open reading frames (ORF) that encode structural, non-structural, and accessory proteins. One of the encoded structural proteins, the spike (S) protein, is able to attach to the host's angiotensin-converting enzyme 2 (ACE2) receptor, which mediates the entry of the virus. The binding between S protein and ACE2 receptor triggers several events that eventually lead to viral genome replication and the formation of virus particles (Harrison et al., 2020).

Generally, infection of SARS-CoV-2 leads to respiration and systemic indications, such as cough, dyspnea, muscle soreness, and fever (Zheng, 2020). However, these symptoms may progress to pneumonia, multiorgan dysfunction, renal failure, and acute respiratory distress (ARDS) according to the disease's severity (Harrison et al., 2020). Hence, variations in symptoms and the number of deaths are also detected among infected patients in Indonesia (Aisyah et al., 2020).

One of the suspected factors that could explain this variation is the different immunological backgrounds within the Indonesian population. In addition, several studies have also suggested that

live attenuated vaccines (LAVs), such as OPV (Oral Poliovirus Vaccine), BCG (Bacillus Calmette-Guérin), and MMR (Measles, Mumps, and Rubella) vaccines may offer nonspecific protection - trained innate immunity - against other infections, including SARS-CoV-2 (Comunale et al., 2021; Haddad-Boubaker et al., 2021; Malave Sanchez et al., 2021). The nonspecific protection can also be at the adaptive immune level where T cells that are primed by LAV confer protection against SARS-CoV-2, which was suggested by Tomita et al. (2020), which shows the cross-reactive epitopes presence of SARS-CoV-2 and BCG.

Cross-reactivity is able to reveal the extent of unexpected reactivity of dissimilar antigens to the target, resulting in immune responses (Tomita et al., 2020). A study by Grifoni et al. (2020) showed that cross-reactive T cell epitopes are able to induce responses of adaptive immunity to viruses, such as SARS-CoV-2. Thus, the cross-reactivity between SARS-CoV-2 and LAVs may bring potential aftermath on the overall course of the ongoing pandemic (Kissler et al., 2020). This motivates us to look for the presence of cross-reactive epitopes in other LAVs in the context of the adaptive immune responses of the Indonesian population.

In this research project, the Wuhan-Hu-1 isolate is the protein of interest that will be analyzed to predict T cell epitopes specific to the Indonesian population. These epitopes will be examined through immunoinformatics techniques to investigate the cross-reactivity against the LAVs (OPV, BCG, and MMR). The identified cross-reactive epitopes will be further inspected for their population coverage in Indonesia. These cross-reactive patterns may also serve as potential epitopes for developing SARS-CoV-2 vaccine construct specific to the Indonesian population. Hence, the primary objective of this study is to discover potential cross-reactive T cell epitopes between SARS-CoV2 and LAVs (OPV, BCG, and MMR) specific for the Indonesian population, using a combination of immunoinformatics approaches.