**Limitations of One-Hot Encoding**

1.Variable Length: The lengths of CDR3 sequences can vary significantly, which complicates the direct application of one-hot encoding since it requires inputs to have uniform dimensions.

2.Dimensionality Explosion: Even if a way is found to accommodate variable lengths, one-hot encoding of CDR3 sequences can lead to extremely high dimensions, especially considering the vast number of possible amino acid combinations.

3.Sparsity: The matrices produced by one-hot encoding are highly sparse due to most encoding positions being 0, which can impact the efficiency and effectiveness of certain machine learning algorithms.

**Overcoming Methods**

1.Fixed length truncation/padding: To address the issue of length variation, all CDR3 sequences can be truncated or padded to the same length. This method is simple and direct, but it may lead to information loss or introduce noise.

2.Variable length encoding scheme: Develop an encoding method for handling variable length sequences, such as using recursive neural networks (RNNs) or attention mechanisms, which can naturally handle inputs of different lengths.

3.Sequence embedding: Using deep learning techniques to map CDR3 sequences to a low dimensional continuous vector space. Through training, these embeddings can capture the complex relationships between sequences and the physicochemical properties of amino acids.