The exponential growth of genomic data calls for novel space-efficient algorithms for compression and search. State-of-the-art approaches often rely on tokenization of the data into k-mers, which are substrings of a fixed length. The popularity of k-mer based methods has led to the development of compact textual k-mer set representations, however, these rely on structural assumptions about the data which may not hold in practice. In this thesis, we demonstrate that all these representations can be viewed as superstrings of the k-mers, and as such can be generalized into a unified framework that we call the masked superstrings of k-mers. We provide two different greedy heuristics for their computation and implement them in a tool called KmerCamel. We further demonstrate that masked superstrings can serve as a building block of a novel, simple k-mer set index which we call FMS-index. Additionally, if masked superstrings further integrate a demasking function f, the resulting f-masked superstrings framework allows for seamless set operations with k-mers. We experimentally evaluate the performance of masked superstrings, as well as of our FMS-index implementation, FMSI, and show that masked superstrings achieve better compression in situations where the previous methods were far from optima. Furthermore, we demostrate that using FMS-index leads to memory savings compared to state-of-the-art indexing methods. Overall, our results demonstrate the usefulness of masked superstrings as a unified theoretical framework as well as their potential in designing data structures for k-mers.