

Cytochrome P450 enzymes are hemoproteins showing extraordinary diversity in the reactions they catalyze. We developed a database containing all the needed data to provide a comprehensive data source on reactions catalyzed by cytochrome P450 enzymes. This data mainly includes information about the substrates, products of characterized reactions, and the sequence of these enzymes. The database was developed by collecting data from reliable protein and reaction databases like UniProt and RHEA. The work presents an in-depth analysis of the created database of reactions catalyzed by cytochrome P450 enzymes. This database can be utilized for future machine learning approaches to predict the function of uncharacterized cytochrome P450s.