




Task 1: EC Number Prediction

Instruction	<p>What is the EC number of the input sequence?</p>
Input	<p>MSLLAYTNLLLQNGRIFRYYYKKANIKKFIKKIILDLKSTPSEASVSRQTFLSTGLNSVKNAVQLQARKLLINNVLERVTPTLNSDLKKKAAKRLFYGDSAPFFALVGVSLASGS GLLTKDDELEGICWEIREAVSKGKWNDSESENVQLQAANLDELDLGEPIAKGCNAVVYSAKLKNVQSNKLAHQLAVKMMFNVDVESNSTAILKAMYRETVPAMS YFFNQN LFNIENISDFKIRLPPHPNIVRMYS VFADRIPDLQCNKQLYPEALPPRINPEGSGRNMSLFLVMKRYDCTLKEYLRDKTPNMRSSILLLSQLLEAVAHMNIHNISHRDLKSDNILVD LSEGDAYPTIVITDFGCCLCDKQNGLVIPYRSEDQDKGGNRALMAPEIANAKPGTFSWLNYYKSDLWAVGAIAYEIFNIDNPFYDKTMKLLSKSYKEEDLPELPDTIPFIIRNLVS NMLSRSTNKRLDCDVAA TVAQLYLWAPSSWLKENYTL PNSNEIIQWLLCLSSK VLCERDITARNKTNMTSESVSKAQYKGRRSLPEYELIASFLRRVRLHLVRKGLKWIQELHI YN</p>
Ground Truth	<p>2.7.11.1</p>
OPI-tuned model	<p>2.7.11.1</p>
 GPT-4o	<p>The input sequence provided is a protein sequence, not an RNA sequence. To determine the EC number (Enzyme Commission number), which classifies enzymes based on the chemical reactions they catalyze, you'd need to identify the enzyme based on the sequence. Here's a general approach to identify the enzyme: <b>1.Sequence Alignment:</b> Use the sequence to perform a BLAST search against a protein database to identify similar sequences and determine if it's an enzyme with a known EC number. <b>2.Annotation Databases:</b> Use databases like UniProt or InterPro to annotate the sequence and find if it matches any known enzymes with an EC number. Would you like me to help you with one of these approaches, or do you have any specific databases or tools in mind?</p>
 Llama-3.1-8B-Instruct	<p>I'd be happy to help! The EC number of the input sequence is not directly provided, as EC numbers are assigned to enzyme classification, and the input sequence does not appear to be an enzyme. EC numbers are typically used to classify enzymes based on their catalytic activity and reaction mechanism. However, if you're interested in identifying the protein function or classifying the input sequence, I can assist you with that. Please let me know what specific information you're looking for, and I'll do my best to help!</p>
 Claude 3.5 Sonnet	<p>To determine the EC number of an input protein sequence, I would typically need to perform a sequence similarity search against databases of known enzymes or use specialized tools for enzyme classification. However, I don't have the capability to directly access such databases or run those tools. From just examining the sequence itself, I can't definitively state its EC number. EC numbers are assigned based on the enzymatic function of a protein, not solely on its sequence. However, I can make some general observations about the sequence: 1. It appears to be a protein sequence, as it's composed of standard amino acid single-letter codes. 2. The sequence is quite long, containing several hundred amino acids. 3. It contains regions that suggest it might be a kinase, such as the presence of a potential catalytic domain with phrases like "DLKSDNILVDLSEGDAYPTIVITDFG" which is reminiscent of kinase motifs. If this protein is indeed a kinase, it might fall under EC 2.7.- (Transferases transferring phosphorus-containing groups). However, this is just a speculation based on sequence features, and would need to be confirmed through proper functional characterization or bioinformatics analysis. To get the actual EC number, you would need to: 1. Use sequence similarity search tools like BLAST against enzyme databases. 2. Use specialized enzyme prediction tools. 3. Look up literature about this specific protein if it has been previously characterized. If you have access to such tools or additional information about this protein's function, that would help in determining its EC number more accurately.</p>