AR Protein Visualization

Vir Vaidya, Samhitha Sangaraju, Alexander Kim
August 7th, 2024

The Team

Alex

- Morristown, NJ
- Rising Senior at Ohio State
- Computer Science with a Minor in Communication

Samhitha

- Edison, NJ
- Rising Junior at Rutgers
- Computer Science and Data Science with a Minor in Statistics

Vir

- Princeton, NJ
- Rising Senior at Rutgers
- Computer Engineering with minor Computer Science

Motivation

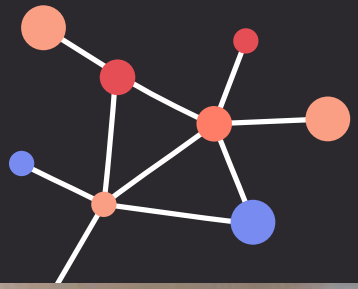
- What is Augmented Reality?
  - Computer-generated image on a user's view of the real world
- Why Augmented Reality?
  - Viewing 3D data on a 2D screen hides many details
  - Immersive and interactive exploration of complex datasets

Project and Technology Overview

Goal: Develop an interactive Augmented Reality (AR) solution to visualize and analyze molecular proteins, enhancing understanding.

- Data Source: Rutgers Protein Data Bank
- Technology: Unity Game Engine with C# scripting, Microsoft Hololens 2

RCSB PDB
PROTEIN DATA BANK



Program Breakdown

Input code: "*6PHQ*"

RCSB PDB PROTEIN DATA BANK

OR

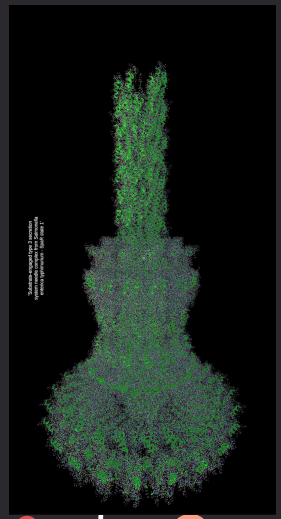
PDB

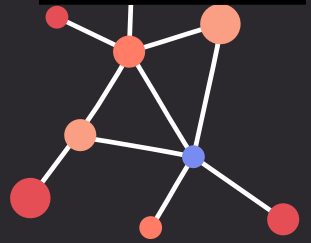
CIF

CSV

→ Color Atoms

→ Combine Meshes

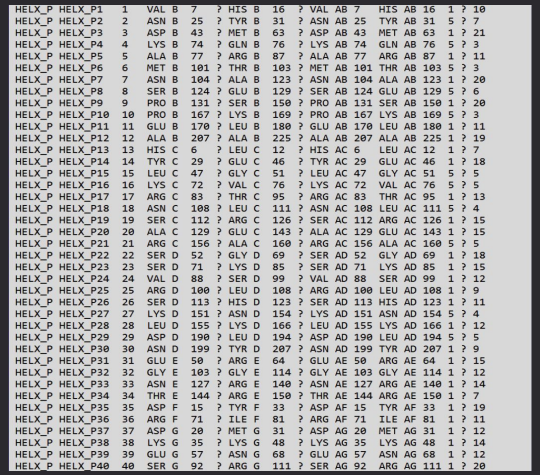




Web Scraping and CSV

- Downloads file from database
- Filters important information
- Autonomous

Original File:



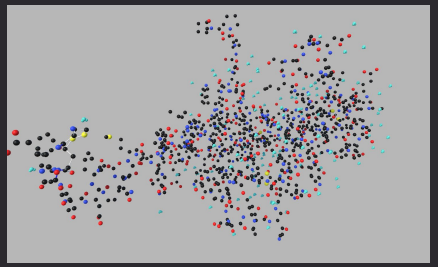
```
HELX_P HELX_P1   1   VAL B  7   ? HIS B  16  ? VAL AB 7   HIS AB 16  1 ? 10
HELX_P HELX_P2   2   ASN B  25  ? TYR B  31  ? ASN AB 25  TYR AB 31  5 ? 7
HELX_P HELX_P3   3   ASP B  43  ? MET B  63  ? ASP AB 43  MET AB 63  1 ? 21
HELX_P HELX_P4   4   LYS B  74  ? GLN B  76  ? LYS AB 74  GLN AB 76  5 ? 3
HELX_P HELX_P5   5   ALA B  77  ? ARG B  87  ? ALA AB 77  ARG AB 87  1 ? 11
HELX_P HELX_P6   6   MET B  101 ? THR B  103 ? MET AB 101 THR AB 103 5 ? 3
HELX_P HELX_P7   7   ASN B  104 ? ALA B  123 ? ASN AB 104 ALA AB 123 1 ? 20
HELX_P HELX_P8   8   SER B  124 ? GLU B  129 ? SER AB 124 GLU AB 129 5 ? 6
HELX_P HELX_P9   9   PRO B  131 ? SER B  150 ? PRO AB 131 SER AB 150 1 ? 20
HELX_P HELX_P10  10  PRO B  167 ? LYS B  169 ? PRO AB 167 LYS AB 169 5 ? 3
HELX_P HELX_P11  11  GLU B  170 ? LEU B  180 ? GLU AB 170 LEU AB 180 1 ? 11
HELX_P HELX_P12  12  ALA B  207 ? ALA B  225 ? ALA AB 207 ALA AB 225 1 ? 19
HELX_P HELX_P13  13  HIS C  6   ? LEU C  12  ? HIS AC 6   LEU AC 12  1 ? 7
HELX_P HELX_P14  14  TYR C  29  ? GLU C  46  ? TYR AC 29  GLU AC 46  1 ? 18
HELX_P HELX_P15  15  LEU C  47  ? GLY C  51  ? LEU AC 47  GLY AC 51  5 ? 5
HELX_P HELX_P16  16  LYS C  72  ? VAL C  76  ? LYS AC 72  VAL AC 76  5 ? 5
HELX_P HELX_P17  17  ARG C  83  ? THR C  95  ? ARG AC 83  THR AC 95  1 ? 13
HELX_P HELX_P18  18  ASN C  108 ? LEU C  111 ? ASN AC 108 LEU AC 111 5 ? 4
HELX_P HELX_P19  19  SER C  112 ? ARG C  126 ? SER AC 112 ARG AC 126 1 ? 15
HELX_P HELX_P20  20  ALA C  129 ? GLU C  143 ? ALA AC 129 GLU AC 143 1 ? 15
HELX_P HELX_P21  21  ARG C  156 ? ALA C  160 ? ARG AC 156 ALA AC 160 5 ? 5
HELX_P HELX_P22  22  SER D  52  ? GLY D  69  ? SER AD 52  GLY AD 69  1 ? 18
HELX_P HELX_P23  23  SER D  71  ? LYS D  85  ? SER AD 71  LYS AD 85  1 ? 15
HELX_P HELX_P24  24  VAL D  88  ? SER D  99  ? VAL AD 88  SER AD 99  1 ? 12
HELX_P HELX_P25  25  ARG D  100 ? LEU D  108 ? ARG AD 100 LEU AD 108 1 ? 9
HELX_P HELX_P26  26  SER D  113 ? HIS D  123 ? SER AD 113 HIS AD 123 1 ? 11
HELX_P HELX_P27  27  LYS D  151 ? ASN D  154 ? LYS AD 151 ASN AD 154 5 ? 4
HELX_P HELX_P28  28  LEU D  155 ? LYS D  166 ? LEU AD 155 LYS AD 166 1 ? 12
HELX_P HELX_P29  29  ASP D  190 ? LEU D  194 ? ASP AD 190 LEU AD 194 5 ? 5
HELX_P HELX_P30  30  ASN D  199 ? TYR D  207 ? ASN AD 199 TYR AD 207 1 ? 9
HELX_P HELX_P31  31  GLU E  50  ? ARG E  64  ? GLU AE 50  ARG AE 64  1 ? 15
HELX_P HELX_P32  32  GLY E  103 ? GLY E  114 ? GLY AE 103 GLY AE 114 1 ? 12
HELX_P HELX_P33  33  ASN E  127 ? ARG E  140 ? ASN AE 127 ARG AE 140 1 ? 14
HELX_P HELX_P34  34  THR E  144 ? ARG E  150 ? THR AE 144 ARG AE 150 1 ? 7
HELX_P HELX_P35  35  ASP F  15  ? TYR F  33  ? ASP AF 15  TYR AF 33  1 ? 19
HELX_P HELX_P36  36  ARG F  71  ? ILE F  81  ? ARG AF 71  ILE AF 81  1 ? 11
HELX_P HELX_P37  37  ASP G  20  ? MET G  31  ? ASP AG 20  MET AG 31  1 ? 12
HELX_P HELX_P38  38  LYS G  35  ? LYS G  48  ? LYS AG 35  LYS AG 48  1 ? 14
HELX_P HELX_P39  39  GLU G  57  ? ASN G  68  ? GLU AG 57  ASN AG 68  1 ? 12
HELX_P HELX_P40  40  SER G  92  ? ARG G  111 ? SER AG 92  ARG AG 111 1 ? 20
```

CSV File:

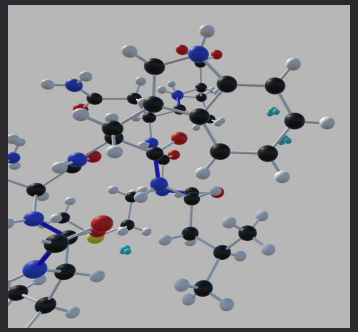
HELIX	7	16 AB
HELIX	25	31 AB
HELIX	43	63 AB
HELIX	74	76 AB
HELIX	77	87 AB
HELIX	101	103 AB
HELIX	104	123 AB
HELIX	124	129 AB
HELIX	131	150 AB
HELIX	167	169 AB
HELIX	170	180 AB
HELIX	207	225 AB
HELIX	6	12 AC
HELIX	29	46 AC
HELIX	47	51 AC
HELIX	72	76 AC
HELIX	83	95 AC
HELIX	108	111 AC
HELIX	112	126 AC
HELIX	129	143 AC
HELIX	156	160 AC
HELIX	52	69 AD

Generation and Interaction

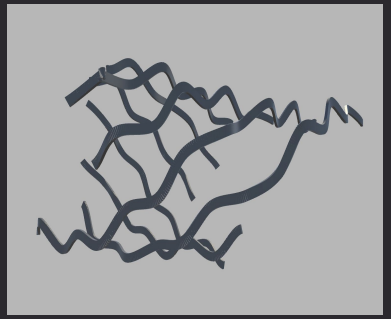
- Protein features from CSV
- Colors and Scales atoms
- Interaction Scripts attached



Atoms

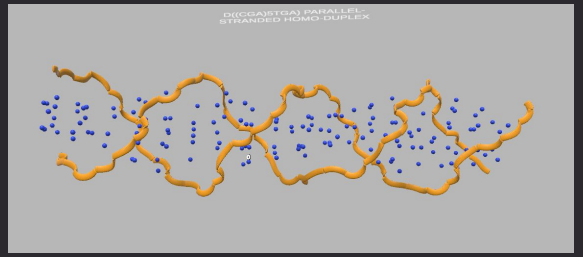


Bonds



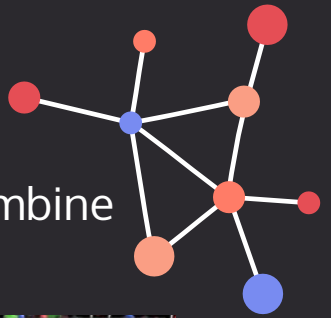
Beta Sheets

Helices

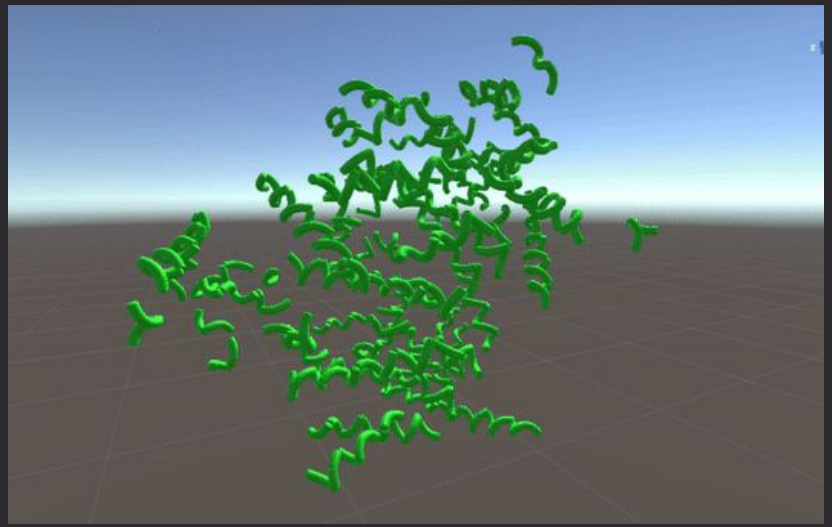


RNA

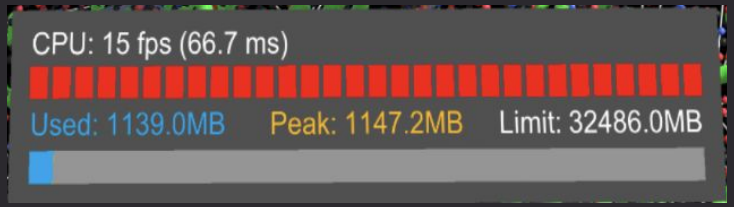
Challenges and Optimization Efforts



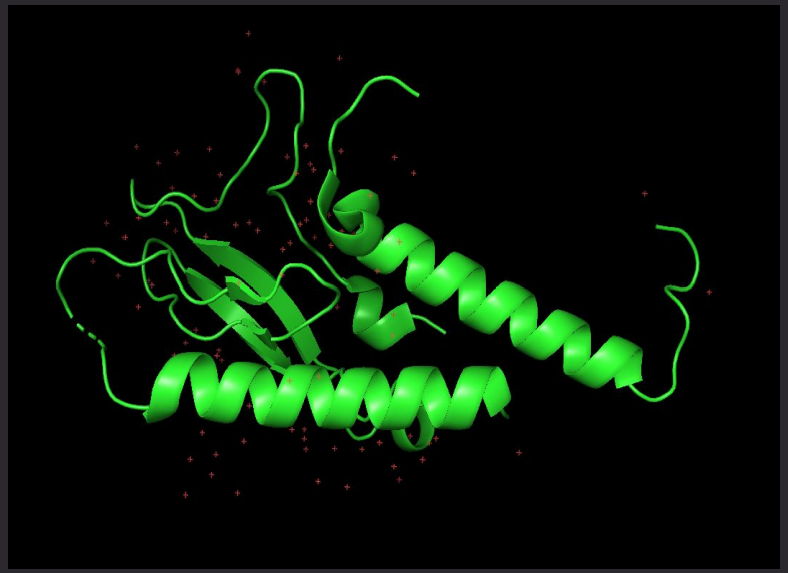
Jagged curved structures → use splines to smooth out



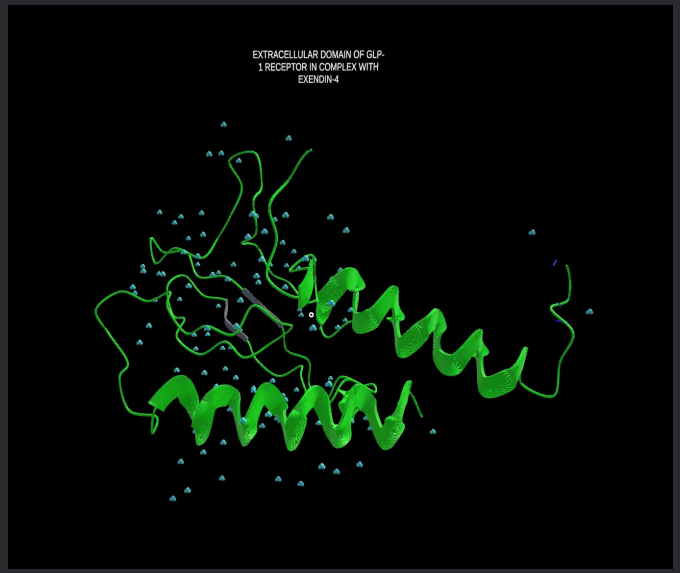
Slow model rendering → combine meshes of similar objects



Results

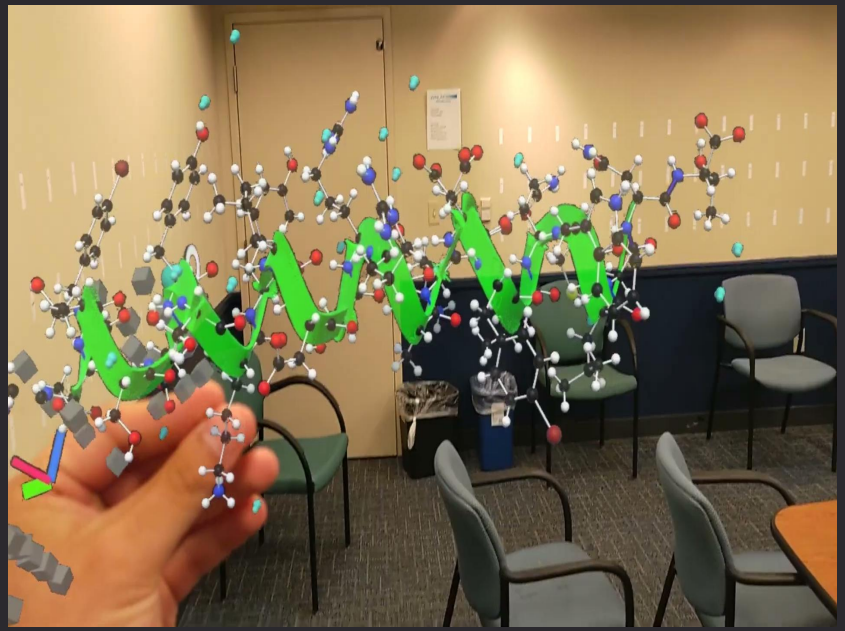


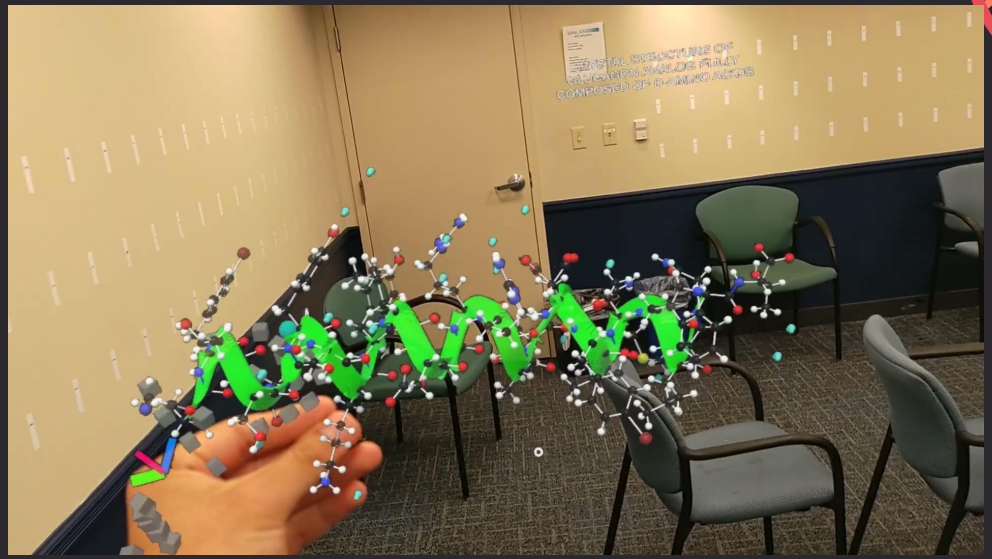
5OTT PyMol Model



5OTT Model in Unity

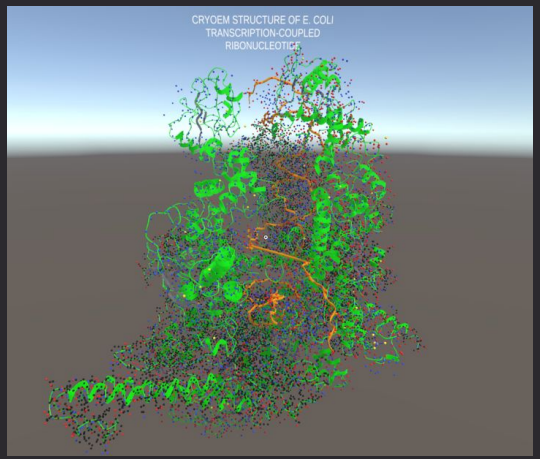
Voice Command Feature





Future Plans

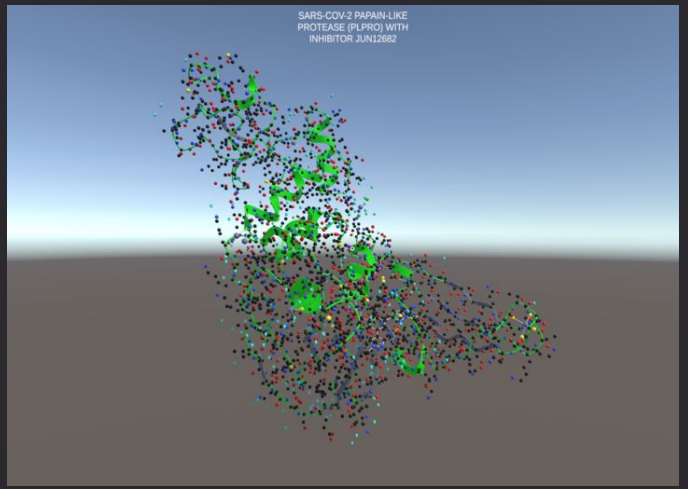
- Greater detail in RNA and Beta-Sheets implementation
- Optimized rendering for large proteins
- Faster 3D model generation and loading time of proteins
- Deploy visualizations to a mobile application
- Feedback from professionals



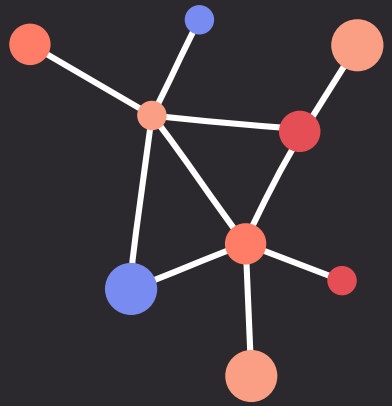
CryoEM Structure of E.Coli

Potential Project Applications

- Interactive Learning Experiences
- Drug Discovery and Manipulation
- Medical Training
- Public Health Awareness



SARS-CoV-2



This work was supported in part by RU-NB CSI Scarlet Vision with assistance from the RCSB PDB team, Ivan Seskar, Jennifer Shane, and Professors Wade Trappe and Ashley Guo.

Thank You!