**0.2 pts** a) Explain both the user and the model inputs of AF models (AF2/AF3 does not matter).

0.15pts protein sequence ⇒ user input
model inputs → MSA : conserved motifs etc.
→ pair representations ; relationship of residues
(relative position information)

0.05pts user input → prot. sequence
if preferred and available, protein structure as a template

b) How do AF models (AF2/AF3 does not matter) flow between MSA representations and pair representations? Explain the vectorial process used to flow information between two parallel processes.

**0.2 pts** using outer product, instead of sum of the outer products, (0.1pts)
takes average of the outer products to flow information
(0.0 pts) from MSA to pair representations. → vectorial operation
(column attention)
(row attention)

c) What is the IPA (invariant point attention) module used by AF models, how does that solve the issue about knowing the exact coordinates?

**0.2pts** instead of knowing (predicting) the exact coord., which depends on
identification of a reference point (origin), compute **relative**
positions of atoms, which is independent of the coordinate
system.

d) Why do AF models use triangle representation? How does that contribute to the structure prediction?
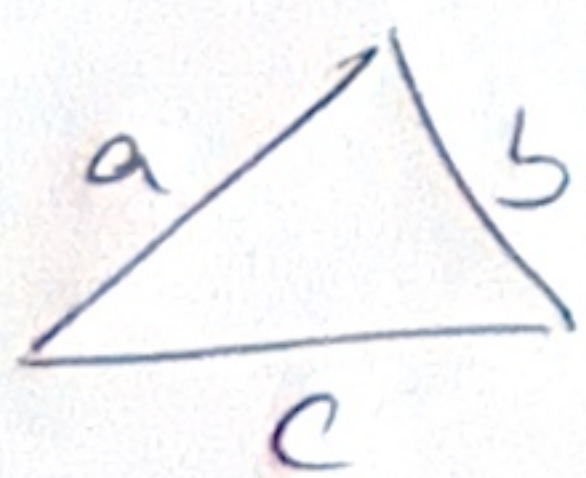
**0.2 pts** → geometric relationships among residues
it significantly constraints search


c < a+b
triangle equation

space and show how three residues
constrain each other's positions in
3D space.

ĉ ĵ – close
î k̂ – close

ĵ ĉ must
be close

e) Explain two differences of AF2 and AF3 in their workflows.

**0.2 pts** model → AF2: input prot. seq,
inputs: → AF3: prot. seq. / DNA/RNA seq. / metal ions /
ligands / lipids
diversity in input
representation

→ (IPA)
structural → AF2: take learned rep. and produce atomic coord.
models: (backbone +)
→ AF3: diffusion module → raw coord. (side chains)
at once