

# Calculating Protein Ionization and Residue pK

**Required functionality and modules:** Discovery Studio Client

**Required data files:** 2lzt.pdb.

**Time:** 15 minutes.

## Introduction

The calculations of the protein ionization and residue pKa's in Discovery Studio are based on a fast and accurate computational approach to pH-dependent electrostatic effects in protein molecules. The protocol also includes a CHARMM-based algorithm to construct and optimize the hydrogen coordinates at a given pH. The detailed description of the methods used can be found in [\[Spassov and Yan, 2008\]](#).

In this tutorial you will learn how to calculate the titration characteristics of a protein molecule and how to predict the pK values of the individual sites of titration on a lysozyme structure. All calculations are carried out using [Protein Ionization and Residue pK](#).

The modeling of the protein ionization is important because the pH-dependent changes in the protonation states of ionizable amino-acid residues can affect almost all molecular mechanisms related to protein function and stability. The Protein Ionization protocol provides an easy and fast way to calculate the complex ionization of the molecule, and at a given pH, to re-protonate and optimize the positions of all hydrogen atoms according to the predicted pKa values.

The tutorial covers:

- [Preparing the molecule and running the calculations](#)
- [Analyzing the predicted pK values of the individual amino-acid residues](#)
- [Analyzing the protonation of individual residues as a function of pH](#)
- [Analyzing the total charge of molecule as a function of pH](#)
- [Analyzing the pH-dependence of protein stability](#)
- [Analyzing the hydrogen structure at given pH](#)

## Preparing the molecule and running the calculations

In the **Files Explorer**, open **Samples | Tutorials | 2lzt.pdb**. If the Hierarchy View is not displayed, choose **View | Hierarchy** from the menu bar.

Before you run the calculations, we recommend that you remove the water molecules and all non-specific ligands and ions from the structure.

In the Hierarchy View, select all water and NO3 molecules.

Press CTRL+X to remove them.

Before running the protocol, apply the CHARMM forcefield to the protein structure.

In the **Tools Explorer**, open the **Simulation** layout.

In the Simulate Structures tool panel, select the **CHARMM** forcefield and click **Apply Forcefield**.

Now open the protocol and modify the parameters.

In the **Protocols Explorer**, expand the **Electrostatics** folder and double-click **Calculate Protein Ionization and Residue pK**.

In the Parameters Explorer, click the **Input Typed Protein Molecule** parameter and choose **2lzt:2lzt** from the dropdown list.

Set the value of **Dielectric Constant** to **11**.

Set **Protonate Protein** to **True**.

Set **Protonate Protein/At pH** to **5.5** to observe the structure of the protonated state of Glu 35 from the active site of enzyme ( the experimental pKa of Glu35 is 6.1). Leave all other parameters with the default values.

In the Protocols toolbar, click **Run**  and wait for the job to complete.

The Job Completed dialog displays when the job is complete.

In the **Jobs Explorer**, double-click the completed job .

This opens a Report.htm file in an Html Window.

## Analyzing the predicted pK values of the individual amino-acid residues

The calculation of the ionization of individual residues is important because it enables the derivation of all other pH-dependent characteristics of a protein molecule. A pK value of an amino-acid reported in this protocol is defined as the pH value at which the titratable group is half protonated.

From **Report.html**, click **Output Files/Annotated Protein Molecule/2lzt\_out.dsv**.

The output structure appears in the Molecule Window. You can now inspect the predicted pK values of the individual acidic and basic groups.

Expand the **Data Table** and open the *AminoAcid* tab.

The predicted pK values are stored in three columns as *Predicted C-term pK*, *Predicted N-term pK*, and *Predicted pK* for the C-terminal, N-terminal, and side-chain pK values respectively. The pK values are also stored in a simple text file.

From **Report.html**, click **Output Files/Predicted pK Values/2lzt.pK**.

Analyze the predicted pK values. In particular, notice that some of the acidic residues have unusually high or low pK values when compared to the values of ~4. and ~4.5 for small model peptides Asp and Glu respectively. Note the high value of 5.73 predicted for Glu 35.

The active center of the enzyme is consistent with the function of this residue as a proton donor and it is quite close to the experimental value of 6.1 - 6.2.

## Analyzing the protonation of individual residues as a function of pH

To analyze the protein ionization of individual residues as well as the titration characteristics of entire molecule, open a csv file with the stored results.

In the **Report.html** window, click **Output Files/Titration Curves/2lzt.csv**.

A window opens with the titration data. Now create a plot of the fractional protonation of the titratable residues from the active center Glu 35 and Asp 52.

In the menu bar, choose **Chart | Point Plot**.

This opens a Select Plot Axis dialog.

For the X axis, choose **pH**. For the Y axis, choose **A:GLU35 and A:ASP52**.

The pH-dependence of the fractional protonation of the selected residues appears as two sigmoidal curves.

## Analyzing the total charge of molecule as a function of pH

To analyze the changes in the total electric charge of molecule with pH of the solution, plot the predicted titration curve of a lysozyme.

In the menu bar, choose **Chart | Point Plot**.

For the X axis, choose **pH**. For the Y axis, choose **Total Charge**.

The pH-dependence of integral ionization of titratable residues appears as a plot of Total Charge versus pH. Notice the total charge is 0 where pH = 11.4, which is in agreement with the experimental value of isoelectric point, pI = 11.2.

## Analyzing the pH-dependence of protein stability

The protocol provides an approximate estimation of the pH-dependent contribution to protein free-energy. You can observe the changes in folding energy by creating a plot of the relative folding energy as a function of pH.

In the menu bar, choose **Chart | Point Plot**.

Choose **pH** for the X axis and **Relative Folding Energy** for the Y axis.

The plot of the Relative Folding Energy versus pH shows an optimum of protein stability at pH = 4.4.

## Analyzing the hydrogen structure at given pH

When the parameter Protonate Molecule is set to **True**, an optimization of the hydrogen network is carried out at the selected pH value. This optimization includes:

- Assignment of hydrogen atoms of the titratable groups according the calculated pK values
- Search for optimal tautomeric forms of histidine residues
- Flipping of amide groups of Asn and Gln and protonated Asp and Glu
- Final energy optimization of all hydrogen positions

You can easily analyze the differences between the hydrogen coordinates in input and output structures.

In the **Report.html** window, click **Input Files/2lzt.dsv**.

This opens a new Molecule Window with the input 2lzt.dsv structure.

Drag the tab for the previously open **2lzt\_out Molecule Window** to the bottom of the viewing area to split the view with the 2lzt Molecule Window.

Inspect the structures of the following residues: Glu35, His15, Gln 41, Gln 50, and Asn 106.

Notice that in the output structure, the protonation state of Glu35 is changed from deprotonated to protonated, the tautomeric state of His 15 is changed, and and the amide groups of Gln 41, Gln 50, and Asn 106 are flipped. You can superimpose the structures to observe many differences between the hydrogen coordinates.

### Further information

[Electrostatics](#)

[Protein Ionization and Residue pK](#).