

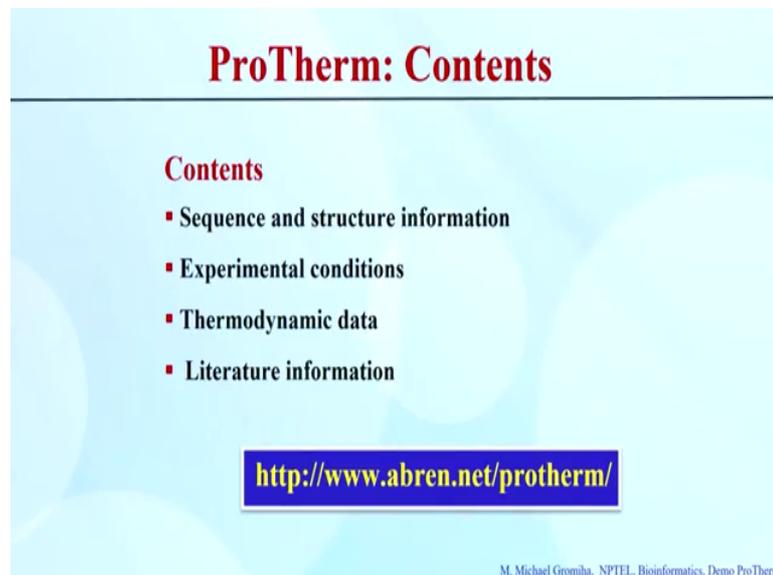
Bioinformatics
Dr. M. Michael Gromiha
Department of Biotechnology
Indian Institute of Technology, Madras

Lecture – 49

Demonstration on ProTherm: Thermodynamic database for proteins and mutants

Demonstration on ProTherm is Thermodynamic Database for proteins and mutants.

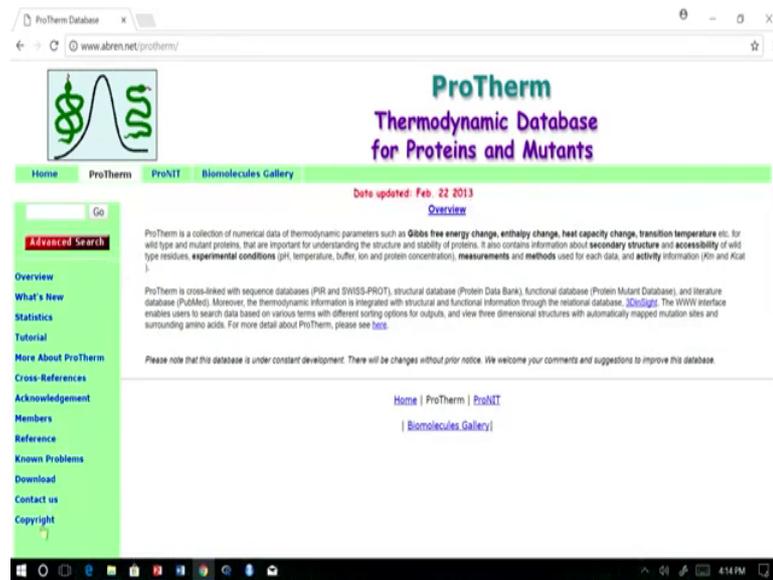
(Refer Slide Time: 00:28)



The slide features a light blue background with a darker blue header bar. The title 'ProTherm: Contents' is written in a bold, red, serif font. Below the title, the word 'Contents' is written in a smaller, red, serif font. A bulleted list follows, with each item preceded by a small red square. The items are: 'Sequence and structure information', 'Experimental conditions', 'Thermodynamic data', and 'Literature information'. At the bottom center, a dark blue rectangular box contains the URL 'http://www.abren.net/protherm/' in a yellow, sans-serif font. In the bottom right corner, there is a small, faint text credit: 'M. Michael Gromiha, NPTEL, Bioinformatics, Demo ProTherm'.

As we discussed earlier ProTherm contents information on sequence information, experimental conditions thermodynamic data and so, literature information for different proteins, as well as their mutants. The major aspect of ProTherm is thermodynamic data and supplemented with all the other information, we can access ProTherm in this website, abren dot net ProTherm and, explain the utilities and how to retrieve data from ProTherm.

(Refer Slide Time: 00:57)



Oh this is website for ProTherm. So, here you can see the links with the other databases like pronate or biomolecules gallery, and left side, you can see the major aspects of ProTherm, you can see the overview. So, what the major contents of this a database and, what we can obtain from this database and, here we update the, what is new and the statistics page and we give few tutorials.

(Refer Slide Time: 01:25)



So, here you can see details how to obtain data, or how to retrieve data from ProTherm database and, here you can give the more details about the ProTherm all right.

(Refer Slide Time: 01:36)

The screenshot shows the ProTherm Thermodynamic Database homepage. The browser address bar displays "www.abren.net/prothem/prothem_intro.php". The page features a logo with a green protein structure and a graph. The main heading is "ProTherm Thermodynamic Database for Proteins and Mutants". A navigation menu includes "Home", "ProTherm", "ProNT", and "Biomolecules Gallery". A "Data updated: Feb. 22 2013" notice is present. A search bar with a "Go" button and an "Advanced Search" link is visible. A sidebar on the left contains a list of navigation options: Overview, What's New, Statistics, Tutorial, More About ProTherm, Cross-References, Acknowledgement, Members, Reference, Known Problems, Download, Contact us, and Copyright. The main content area includes a "More About ProTherm" link and a paragraph of introductory text about the database's history and goals.

So, now, here is a cross references.

(Refer Slide Time: 01:39)

The screenshot shows the ProTherm Cross-References page. The browser address bar displays "www.abren.net/prothem/prothem_crossreference.php". The page features the same logo and navigation menu as the homepage. A "Data updated: Feb. 22 2013" notice is present. A "Cross-References from PDB PIR and SWISS-PROT to ProTherm" link is visible. The main content area includes a paragraph of introductory text about the correspondence between database entries. A list of four download links is provided: 1. PDB to ProTherm (100%), 2. PDB to ProTherm (95% and higher), 3. PDB to ProTherm Download, and 4. SWISS-PROT to ProTherm Download. A "Direct linking from PDB to ProTherm entries" section explains how to use the URL "http://www.abren.net/prothem/prothem_pdburl.php?ENTR****" where "****" corresponds to a 4-letter PDB ID. The sidebar and navigation menu are identical to the homepage screenshot.

You can see the PDB map to the ProTherm and Swiss proto ProTherm and so, on.

(Refer Slide Time: 01:48)

The screenshot shows the ProTherm website's 'References' page. The page title is 'ProTherm Thermodynamic Database for Proteins and Mutants'. It features a navigation menu with 'Home', 'ProTherm', 'ProNT', and 'Biomolecules Gallery'. A search bar is located at the top left. The main content area displays a list of six references, each with a numbered title and author information. The references are:

1. "ProTherm and ProNT, thermodynamic databases for proteins and protein-nucleic acid interactions" Kumar MD, Bava KA, Gromiha MM, Parabakaran P, Kijajima K, Uedara H, Sarai A. (2006) Nucleic Acids Res. 34 D204-6, Database issue
2. "ProTherm, version 4.0, Thermodynamic Database for Proteins and Mutants" K. Abulula Bava, M. Michael Gromiha, H. Uedara, Koji Kijajima and A. Sarai (2004). Nucleic Acids Res. 32, D192-D193, Database issue
3. "ProTherm, Thermodynamic Database for Proteins and Mutants, Developments in Version 3.0" M. Michael Gromiha, H. Uedara, J. An, S. Selvaraj, P. Parabakaran and A. Sarai (2002). Nucleic Acids Res. 30, 301-302
4. "Thermodynamic Databases for Proteins and Protein-Nucleic Acid Interactions" A. Sarai, M.M. Gromiha, J. An, P. Parabakaran, S. Selvaraj, H. Kono, M. Oobatake and H. Uedara (2002). Biopolymers 61, 121-126
5. "ProTherm, version 2.0, thermodynamic Database for Proteins and Mutants" M. Michael Gromiha, J. An, H. Kono, M. Oobatake, H. Uedara, P. Parabakaran and A. Sarai (2000) Nucleic Acids Res. 28,285-288.
6. "ProTherm, Thermodynamic Database for Proteins and Mutants" M. Michael Gromiha, J. An, H. Kono, M. Oobatake, H. Uedara and A. Sarai (1999) Nucleic Acids Res. 27, 286-288.

At the bottom of the page, there is a 'Cite ProTherm' link.

And here we give the references this is the major references and the citations, currently if you look at the different versions of these a papers and website, because more than a 1000 right.

(Refer Slide Time: 02:06)

The screenshot shows the ProTherm website's 'Overview' page. The page title is 'ProTherm Thermodynamic Database for Proteins and Mutants'. It features a navigation menu with 'Home', 'ProTherm', 'ProNT', and 'Biomolecules Gallery'. A search bar is located at the top left. The main content area contains an 'Overview' section with the following text:

ProTherm is a collection of numerical data of thermodynamic parameters such as **Gibbs free energy change**, **enthalpy change**, **heat capacity change**, **transition temperature** etc. for wild type and mutant proteins, that are important for understanding the structure and stability of proteins. It also contains information about **secondary structure** and **accessibility** of wild type residues, **experimental conditions** (pH, temperature, buffer, ion and protein concentration), **measurements and methods** used for each data, and **activity information** (Km and Kcat).

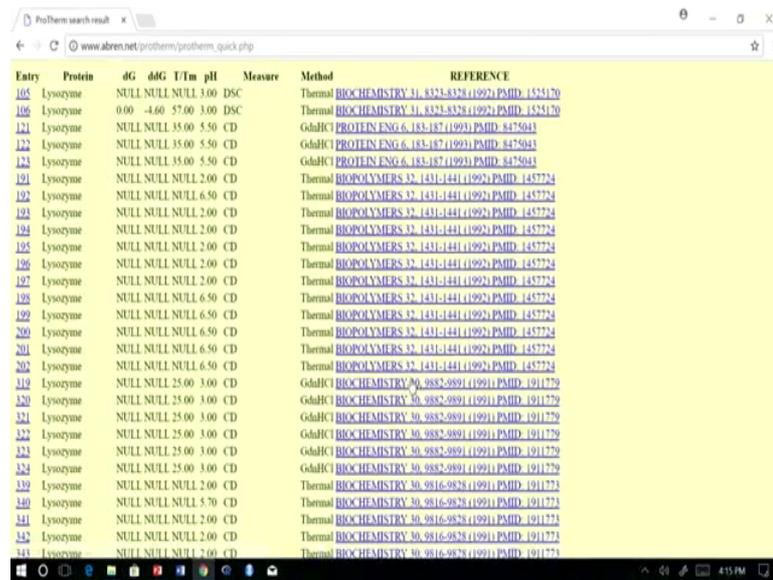
ProTherm is cross-linked with sequence databases (PIR and SWISS-PROT), structural database (Protein Data Bank), functional database (Protein Mutant Database), and literature database (PubMed). Moreover, the thermodynamic information is integrated with structural and functional information through the relational database, **3Dviz**. The WWW interface enables users to search data based on various terms with different sorting options for outputs, and view three dimensional structures with automatically mapped mutation sites and surrounding amino acids. For more detail about ProTherm, please see [here](#).

Please note that this database is under constant development. There will be changes without prior notice. We welcome your comments and suggestions to improve this database.

At the bottom of the page, there are links for 'Home | ProTherm | ProNT | Biomolecules Gallery'.

So, now if you go to the main page right so, so this is the simple search you can give any keywords to search ProTherm for example, if you are interested lysozyme, where this is the protein widely studied, so, can if you click go.

(Refer Slide Time: 02:20)

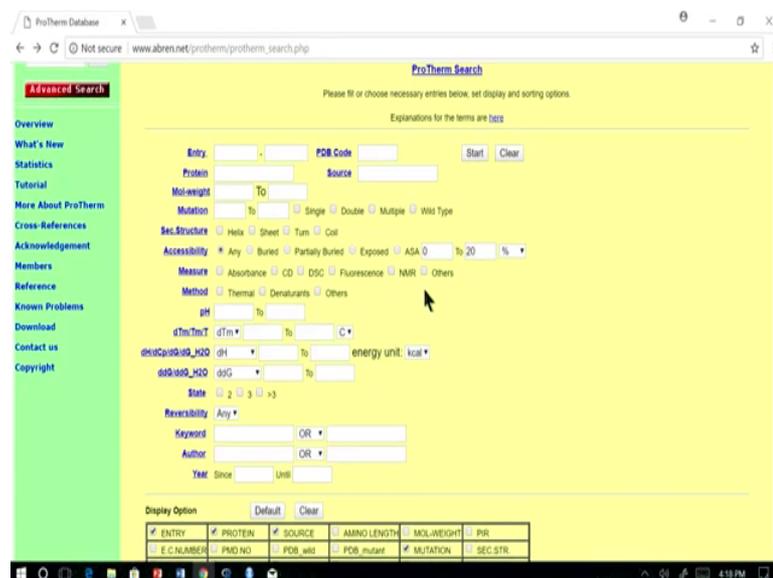


Entry	Protein	dG	ddG	Tm	pH	Measure	Method	REFERENCE
105	Lysozyme	NULL	NULL	NULL	3.00	DSC	Thermal	BIOCHEMISTRY 31, 8323-8328 (1992) PMID: 1525170
106	Lysozyme	0.00	-4.60	57.00	3.00	DSC	Thermal	BIOCHEMISTRY 31, 8323-8328 (1992) PMID: 1525170
121	Lysozyme	NULL	NULL	35.00	5.50	CD	GdHCl	PROTEIN ENG 6, 183-187 (1991) PMID: 8475043
122	Lysozyme	NULL	NULL	35.00	5.50	CD	GdHCl	PROTEIN ENG 6, 183-187 (1991) PMID: 8475043
123	Lysozyme	NULL	NULL	35.00	5.50	CD	GdHCl	PROTEIN ENG 6, 183-187 (1991) PMID: 8475043
191	Lysozyme	NULL	NULL	NULL	2.00	CD	Thermal	BIOPOLYMERS 32, 1431-1441 (1992) PMID: 1457724
192	Lysozyme	NULL	NULL	NULL	6.50	CD	Thermal	BIOPOLYMERS 32, 1431-1441 (1992) PMID: 1457724
193	Lysozyme	NULL	NULL	NULL	2.00	CD	Thermal	BIOPOLYMERS 32, 1431-1441 (1992) PMID: 1457724
194	Lysozyme	NULL	NULL	NULL	2.00	CD	Thermal	BIOPOLYMERS 32, 1431-1441 (1992) PMID: 1457724
195	Lysozyme	NULL	NULL	NULL	2.00	CD	Thermal	BIOPOLYMERS 32, 1431-1441 (1992) PMID: 1457724
196	Lysozyme	NULL	NULL	NULL	2.00	CD	Thermal	BIOPOLYMERS 32, 1431-1441 (1992) PMID: 1457724
197	Lysozyme	NULL	NULL	NULL	2.00	CD	Thermal	BIOPOLYMERS 32, 1431-1441 (1992) PMID: 1457724
198	Lysozyme	NULL	NULL	NULL	6.50	CD	Thermal	BIOPOLYMERS 32, 1431-1441 (1992) PMID: 1457724
199	Lysozyme	NULL	NULL	NULL	6.50	CD	Thermal	BIOPOLYMERS 32, 1431-1441 (1992) PMID: 1457724
200	Lysozyme	NULL	NULL	NULL	6.50	CD	Thermal	BIOPOLYMERS 32, 1431-1441 (1992) PMID: 1457724
201	Lysozyme	NULL	NULL	NULL	6.50	CD	Thermal	BIOPOLYMERS 32, 1431-1441 (1992) PMID: 1457724
202	Lysozyme	NULL	NULL	NULL	6.50	CD	Thermal	BIOPOLYMERS 32, 1431-1441 (1992) PMID: 1457724
119	Lysozyme	NULL	NULL	25.00	3.00	CD	GdHCl	BIOCHEMISTRY 30, 9882-9891 (1991) PMID: 1911729
120	Lysozyme	NULL	NULL	25.00	3.00	CD	GdHCl	BIOCHEMISTRY 30, 9882-9891 (1991) PMID: 1911729
121	Lysozyme	NULL	NULL	25.00	3.00	CD	GdHCl	BIOCHEMISTRY 30, 9882-9891 (1991) PMID: 1911729
122	Lysozyme	NULL	NULL	25.00	3.00	CD	GdHCl	BIOCHEMISTRY 30, 9882-9891 (1991) PMID: 1911729
123	Lysozyme	NULL	NULL	25.00	3.00	CD	GdHCl	BIOCHEMISTRY 30, 9882-9891 (1991) PMID: 1911729
124	Lysozyme	NULL	NULL	25.00	3.00	CD	GdHCl	BIOCHEMISTRY 30, 9882-9891 (1991) PMID: 1911729
139	Lysozyme	NULL	NULL	NULL	2.00	CD	Thermal	BIOCHEMISTRY 30, 9816-9878 (1991) PMID: 1911723
140	Lysozyme	NULL	NULL	NULL	5.70	CD	Thermal	BIOCHEMISTRY 30, 9816-9878 (1991) PMID: 1911723
141	Lysozyme	NULL	NULL	NULL	2.00	CD	Thermal	BIOCHEMISTRY 30, 9816-9878 (1991) PMID: 1911723
142	Lysozyme	NULL	NULL	NULL	2.00	CD	Thermal	BIOCHEMISTRY 30, 9816-9878 (1991) PMID: 1911723
143	Lysozyme	NULL	NULL	NULL	2.00	CD	Thermal	BIOCHEMISTRY 30, 9816-9878 (1991) PMID: 1911723

So, not only you can give all the data regard the lysozyme. So, these are the entry number here, these are protein name and these thermodynamic parameters right, this is the conditions and, which measurement they use to get the data and, the method whether it is thermal or the denaturant and, it can be the complete reference right.

So, now if you are specifically interested on any type of data, then you could advances.

(Refer Slide Time: 02:47)



ProTherm Search

Please fill or choose necessary entries below, set display and sorting options.

Explanations for the terms are here

Entry: PDB Code: Start Clear

Protein: Source:

Mol-weight: To: Single Double Multiple Wild Type

Mutation: To: Single Double Multiple Wild Type

Sec. Structure: Helix Sheet Turn Coil

Accessibility: Any Buried Partially Buried Exposed ASA 0 to 20 %

Measure: Absorbance CD DSC Fluorescence NMR Others

Method: Thermal Denaturants Others

pH: To:

dTm/Tm: To:

deltacp/deltah_H2O: dH: To: energy unit: kcal

deltah/deltah_H2O: To:

State: 2 3 >3

Reversibility: Any

Keyword: OR

Author: OR

Year: Since Until

Display Option: Default Clear

<input checked="" type="checkbox"/> ENTRY	<input checked="" type="checkbox"/> PROTEIN	<input checked="" type="checkbox"/> SOURCE	<input type="checkbox"/> AMINO LENGTH	<input type="checkbox"/> MOL-WEIGHT	<input type="checkbox"/> PIR
<input type="checkbox"/> E.C. NUMBER	<input type="checkbox"/> PDB NO.	<input type="checkbox"/> PDB wild	<input type="checkbox"/> PDB mutant	<input checked="" type="checkbox"/> MUTATION	<input type="checkbox"/> SEC STR.

Right here you have a lot of options available to search, you can search with their entry this is the ProTherm entry. So, this is only specific to a ProTherm that, if you have any

queries on any data, then you can contact the developers using the entry number; it is easy to check the data.

So, you can also search with the PDB code right, then you can use the protein name and the source and, if you know the size of the protein for example, if you have the data on the similar size of proteins for example, a 100 amino acid residues, you can give the molecular weight and, you can give the mutation for example, if you want to see the mutation in from lysine to alanine ok.

So, I can lysine and alanine, what will happened to mutate this one and, here you can choose where a single mutation, or double mutation, or the multiple mutation and if it is single mutation and if you start right.

(Refer Slide Time: 03:35)

Entry	Protein	Source	Mutation	dG_H2O	dG	T	Tm	dHcal	Cm	pH	Measure	Method	REFERENCE
29	Banase	Bacillus amyloliquef	L 14 A	4.29	NULL	25.00	NULL	NULL	1.87	2.20	6.30	Fluorescence Urea	JMOL BIOC 224, 783-804 (1992) PMID: 1569557
109	Staphylococcal nuclease	Staphylococcus aureus	L 25 A	NULL	NULL	NULL	39.90	90.00	NULL	NULL	7.00	DSC	PROTEIN SCI 2, 567-576 (1993) PMID: 8518730
115	Staphylococcal nuclease	Staphylococcus aureus	L 25 A	NULL	NULL	NULL	31.20	63.00	NULL	NULL	5.00	DSC	PROTEIN SCI 2, 567-576 (1993) PMID: 8518730
263	Myoglobin	Sperm whale	L 29 A	NULL	NULL	NULL	67.60	NULL	NULL	NULL	11.00	DSC	BIOCHEMISTRY 32, 12638-12643 (1993) PMID: 8251481
267	Myoglobin	Sperm whale	L 69 A	NULL	NULL	NULL	72.70	NULL	NULL	NULL	11.00	DSC	BIOCHEMISTRY 32, 12638-12643 (1993) PMID: 8251481
219	Staphylococcal nuclease	Staphylococcus aureus	L 7 A	3.88	NULL	20.00	NULL	NULL	0.89	0.60	7.00	Fluorescence GdnHCl	BIOCHEMISTRY 29, 8033-8041 (1990) PMID: 2261461
221	Staphylococcal nuclease	Staphylococcus aureus	L 14 A	3.18	NULL	20.00	NULL	NULL	1.05	0.40	7.00	Fluorescence GdnHCl	BIOCHEMISTRY 29, 8033-8041 (1990) PMID: 2261461
232	Staphylococcal nuclease	Staphylococcus aureus	L 25 A	2.78	NULL	20.00	NULL	NULL	1.12	0.30	7.00	Fluorescence GdnHCl	BIOCHEMISTRY 29, 8033-8041 (1990) PMID: 2261461
235	Staphylococcal nuclease	Staphylococcus aureus	L 36 A	1.98	NULL	20.00	NULL	NULL	1.13	0.20	7.00	Fluorescence GdnHCl	BIOCHEMISTRY 29, 8033-8041 (1990) PMID: 2261461
237	Staphylococcal nuclease	Staphylococcus aureus	L 37 A	3.78	NULL	20.00	NULL	NULL	0.89	0.60	7.00	Fluorescence GdnHCl	BIOCHEMISTRY 29, 8033-8041 (1990) PMID: 2261461
239	Staphylococcal nuclease	Staphylococcus aureus	L 38 A	3.78	NULL	20.00	NULL	NULL	0.94	0.60	7.00	Fluorescence GdnHCl	BIOCHEMISTRY 29, 8033-8041 (1990) PMID: 2261461
431	Staphylococcal nuclease	Staphylococcus aureus	L 89 A	2.88	NULL	20.00	NULL	NULL	1.00	0.40	7.00	Fluorescence GdnHCl	BIOCHEMISTRY 29, 8033-8041 (1990) PMID: 2261461
433	Staphylococcal nuclease	Staphylococcus aureus	L 103 A	0.88	NULL	20.00	NULL	NULL	0.96	0.10	7.00	Fluorescence GdnHCl	BIOCHEMISTRY 29, 8033-8041 (1990) PMID: 2261461

We will get all the lysine to alanine mutations right. So, here we will get the experimental data, I will explain the details how to get this a all this data, then you can search with any secondary structure with the mutation pulse and helix, or sheet or turn or coil and also locations based on the accessibility, you can use either buried, or partially we are exposed these are predefined.

If I specifically interested on any specific range of accessibility surface area you can use here for example, 0 to 20 a angstrom square r 0 to 30 percentage, they give the values both in percentage right and angstrom square right, what say which is a units you gets

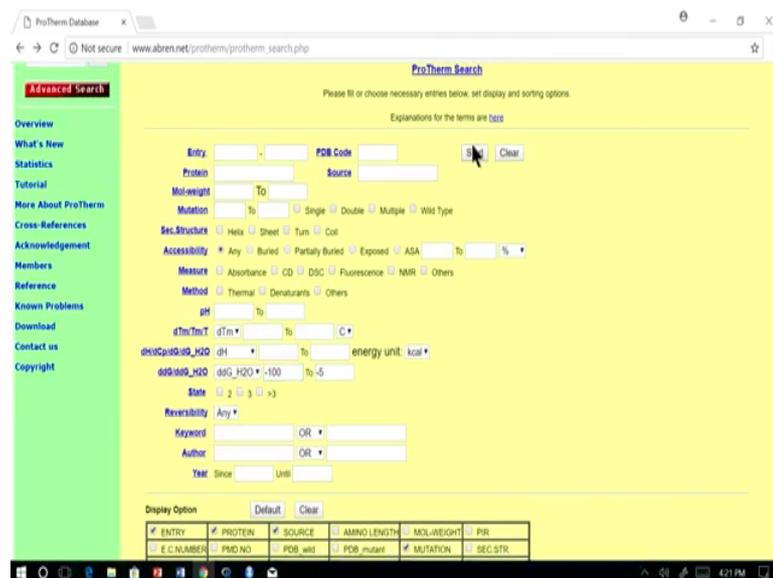
using DSSP, you get I see an X X square, how do we the percentage values. You would do it with the X units, or accessibility and give the percentage and, you can choose any of these options either the directly from the DSSP result, or you can use the percentage you can normalize with the X and X accessibility.

Then we get the measures whether you say CD or DSC or fluorescence these are the major measurements and, the methods whether this is a thermal or the denaturant right for example, if you use the data for the thermal right, we will get a thermal denaturation and the denaturant right. If denaturant thermal means you will get the data for the delta Tm right and, this is the data of the denaturant so, this why it is completely null.

So, if you want to get the denaturants then we will click on denaturant, you will get the data with the denaturants denaturation, these are conditions so, pH and for the denaturants we need to give the temperature right, you can decide temperature.

And also you can search for the different ranges of these experimental data for example, if you are interested in the very extreme stability for example, extremely stabilize, or extremely destabilize right. In this case you can give the delta delta G values as more than minus 10 kilo cal per mole and, you can see the any mutations which are extremely stable or unstable see delta delta G H 2 O right.

(Refer Slide Time: 05:49)



For example you can say you can use it minus 5 all right, then if you see this.

this is more than 5 so, you can see now if the data will be interesting. So, you can see this delta G H 2 O, you can see these are the mutation not many mutations, but still or 32 mutations which are have extremely stable right.

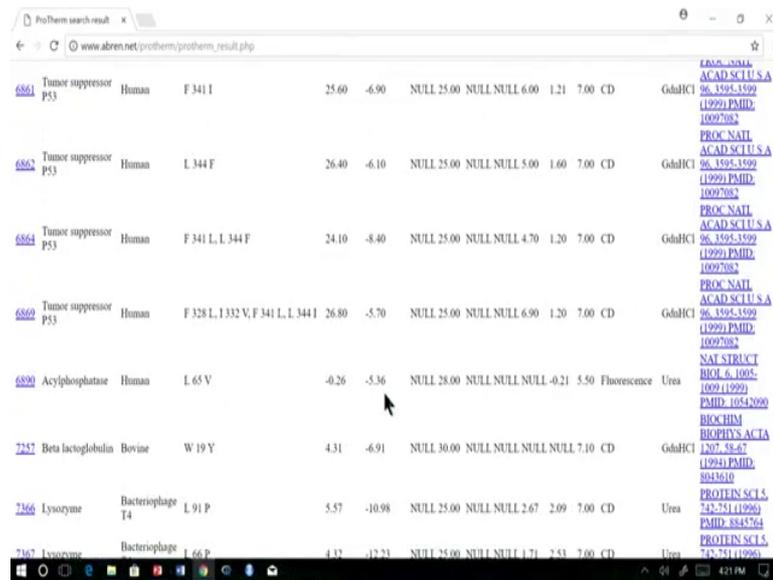
(Refer Slide Time: 07:12)

Entry	Protein	Source	Mutation	dG_H2O	dG_H2O_dG	T	Tm	dHcal	Cm	pH	Measure	Method	REFERENCE	
899	Tryptophan synthase alpha-subu	Escherichia coli	E 49 I	16.80	8.00	18.62	25.00	NULL	NULL	NULL	NULL	7.00	CD	GdbHCl PUBMED:14691476 PUBMED:14691476 PUBMED:14691476
900	Tryptophan synthase alpha-subu	Escherichia coli	E 49 L	15.00	6.20	18.60	25.00	NULL	NULL	NULL	NULL	7.00	CD	GdbHCl PUBMED:14691476 PUBMED:14691476 PUBMED:14691476
918	Tryptophan synthase alpha-subu	Escherichia coli	E 49 I	10.00	5.10	19.21	25.00	NULL	NULL	NULL	NULL	9.00	CD	GdbHCl PUBMED:14691476 PUBMED:14691476 PUBMED:14691476
919	Tryptophan synthase alpha-subu	Escherichia coli	E 49 L	12.20	7.30	19.01	25.00	NULL	NULL	NULL	NULL	9.00	CD	GdbHCl PUBMED:14691476 PUBMED:14691476 PUBMED:14691476
9144	FcpA	Escherichia coli	S 254 C	NULL	5.94	NULL	25.00	NULL	NULL	4.61	1.29	7.00	ESR	GdbHCl PUBMED:14691476 PUBMED:14691476 PUBMED:14691476
9145	FcpA	Escherichia coli	G 252 C	NULL	7.93	NULL	25.00	NULL	NULL	5.28	1.50	7.00	ESR	GdbHCl PUBMED:14691476 PUBMED:14691476 PUBMED:14691476

So, in this particular tryptophan synthase and alba subunit that E 49 I, if you mutated this increases stability of 8 kilo cal per mole. Now, you see this very high right you can see why this happens not for all the mutants, we can see several specific mutations, which are highly stabilizing.

But if we take the highly destabilizing mutants for example, if it is a up to minus 5 right, this is minus 1 100 this is for example, just because we need the complete destabilize once right. If you click on that, then you can see several mutations which are having highly destabilizing.

(Refer Slide Time: 07:58)



Protein search result	Protein Name	Organism	Mutation	ΔG (kcal/mol)	ΔS (cal/mol·K)	ΔH (kcal/mol)	ΔCp (kcal/mol·K)	ΔTm (°C)	CD	Method	References				
6861	Tumor suppressor p53	Human	F 341 I	25.60	-6.90	NULL	25.00	NULL	NULL	6.00	1.21	7.00	CD	GdnHCl	EMBS-35044 ACAD SCITUS A 26, 3595-3599 (1999) PMID: 10927082
6862	Tumor suppressor p53	Human	L 344 F	26.40	-6.10	NULL	25.00	NULL	NULL	5.00	1.60	7.00	CD	GdnHCl	PROC NATL ACAD SCITUS A 26, 3595-3599 (1999) PMID: 10927082
6864	Tumor suppressor p53	Human	F 341 L, L 344 F	24.10	-8.40	NULL	25.00	NULL	NULL	4.70	1.20	7.00	CD	GdnHCl	PROC NATL ACAD SCITUS A 26, 3595-3599 (1999) PMID: 10927082
6869	Tumor suppressor p53	Human	F 328 L, I 332 V, F 341 L, L 344 I	26.80	-5.70	NULL	25.00	NULL	NULL	6.90	1.20	7.00	CD	GdnHCl	PROC NATL ACAD SCITUS A 26, 3595-3599 (1999) PMID: 10927082
6890	Acylphosphatase	Human	L 65 V	-0.26	-5.36	NULL	28.00	NULL	NULL	NULL	-0.21	5.50	Fluorescence	Urea	NAT STRUCT BIOL 6, 1005-1009 (1999) PMID: 10547090
7257	Beta lactoglobulin	Bovine	W 19 Y	4.31	-6.91	NULL	30.00	NULL	NULL	NULL	7.10	7.10	CD	GdnHCl	BIOPHY S ACTA 1207, 58-67 (1994) PMID: 8043610
7366	Lysozyme	Bacteriophage T4	L 91 P	5.57	-10.98	NULL	25.00	NULL	NULL	2.67	2.09	7.00	CD	Urea	PROTEIN SCL 5, 742-751 (1996) PMID: 8845764
7367	Lysozyme	Bacteriophage T4	L 66 P	4.32	-12.33	NULL	25.00	NULL	NULL	1.71	2.53	7.00	CD	Urea	PROTEIN SCL 5, 742-751 (1996) PMID: 8845764

Now, you can take these several mutations and see why these mutations are extremely stable or extremely unstable. So, you can do the analysis and you can see, or you can have any a common features, which you can explain the highly stabled, or highly unstabled mutants and so, on right.

You can use this in absence to check the extremely stabled, as well as extremely a destable mutants these are states as I discussed earlier 2 state means because, the unfolded state and the folded state and if you under mutant states, then you can see 3 or 4 that you can choose these a any states to get your data.

Then you can get reversibility right we can (Refer Time: 08:37) unfolded state and, then you can return back to unfolded state that you say, this is reversibility yes, if not this reversibility is no right like this key words right. So, you can said to the any of the key words as well as the others right and the ears right. So, you can do that.

So, now, we have so many options to search your data right. Now, if you display everything there is a big mess, we do not know what you are searching for and which information do you want. So, in this case we have a display option right. So, you can choose any of these options do you want right, I will explain some of the some of the some examples.

Then we can sorting you can have a sorting options. So, you can sort your data based on the results, the number of entries the default value is 300, but you can use the more number of 3000 30000's these entries. So, you can download all the data all together ok.

Now, I go with these few examples. So, let us see I want to have the stability values for all the single mutants right. So, what to do we got a single mutation right very simple right.

(Refer Slide Time: 09:52)

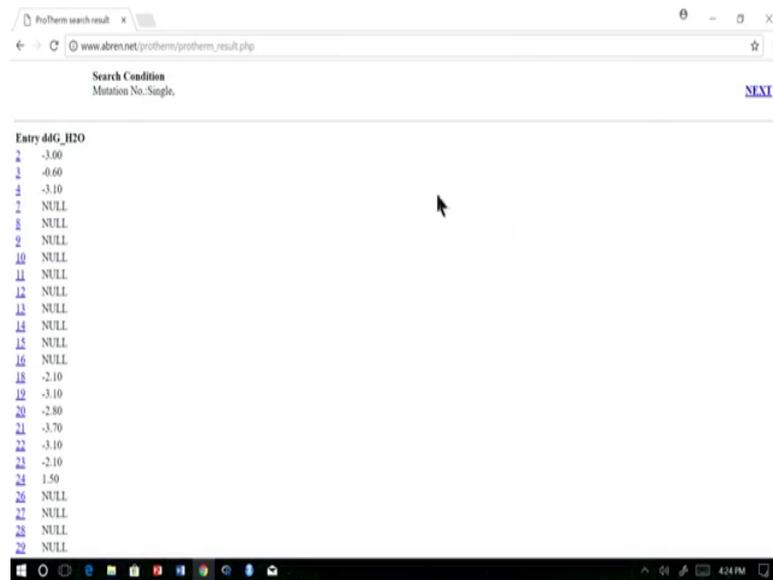
Entry	Protein	Source	Mutation	dG_H2O	dG_T	Tm	dHcal	Cm	pH	Measure	Method	REFERENCE
2	Phospholipase A2	Bovine	H 48 N	6.50	NULL	30.00	NULL	1.20	5.40	8.00	CD	GdaHCl J AM CHEM SOC 115, 8523-8526 (1993) PMID:
3	Phospholipase A2	Bovine	H 48 Q	8.90	NULL	30.00	NULL	1.34	6.60	8.00	CD	GdaHCl J AM CHEM SOC 115, 8523-8526 (1993) PMID:
4	Phospholipase A2	Bovine	H 48 A	6.40	NULL	30.00	NULL	1.02	6.30	8.00	CD	GdaHCl J AM CHEM SOC 115, 8523-8526 (1993) PMID:
7	Ribonuclease HI	Escherichia coli	K 91 R	NULL	NULL	NULL	49.80	NULL	NULL	3.00	CD	Thermal JBIOL CHEM 267, 32014-32017 (1992) PMID: 1331044
8	Ribonuclease HI	Escherichia coli	K 91 R	NULL	NULL	NULL	52.00	NULL	NULL	5.50	CD	Thermal JBIOL CHEM 267, 32014-32017 (1992) PMID: 1331044
9	Ribonuclease HI	Escherichia coli	D 94 E	NULL	NULL	NULL	49.80	NULL	NULL	3.00	CD	Thermal JBIOL CHEM 267, 32014-32017 (1992) PMID: 1331044
10	Ribonuclease HI	Escherichia coli	D 94 E	NULL	NULL	NULL	52.00	NULL	NULL	5.50	CD	Thermal JBIOL CHEM 267, 32014-32017 (1992) PMID: 1331044
11	Ribonuclease HI	Escherichia coli	K 95 G	NULL	NULL	NULL	49.80	NULL	NULL	3.00	CD	Thermal JBIOL CHEM 267, 32014-32017 (1992) PMID: 1331044
12	Ribonuclease HI	Escherichia coli	K 95 G	NULL	NULL	NULL	52.00	NULL	NULL	5.50	CD	Thermal JBIOL CHEM 267, 32014-32017 (1992) PMID: 1331044
13	Ribonuclease HI	Escherichia coli	K 95 A	NULL	NULL	NULL	49.80	NULL	NULL	3.00	CD	Thermal JBIOL CHEM 267, 32014-32017 (1992) PMID: 1331044
14	Ribonuclease HI	Escherichia coli	K 95 A	NULL	NULL	NULL	52.00	NULL	NULL	5.50	CD	Thermal JBIOL CHEM 267, 32014-32017 (1992) PMID: 1331044
15	Ribonuclease HI	Escherichia coli	K 95 N	NULL	NULL	NULL	49.80	NULL	NULL	3.00	CD	Thermal JBIOL CHEM 267, 32014-32017 (1992) PMID: 1331044
16	Ribonuclease HI	Escherichia coli	K 95 N	NULL	NULL	NULL	52.00	NULL	NULL	5.50	CD	Thermal JBIOL CHEM 267, 32014-32017 (1992) PMID: 1331044
18	Interleukin 1 beta	Human	T 9 A	7.00	8.30	25.00	NULL	NULL	1.10	6.50	Fluorescence	GdaHCl JBIOL CHEM 268, 18053-18061 (1993) PMID: 8261352

And if you click start now, it is a big mess you need all these things right So, now, the question is among all the information, which share the you are interested in for example, you have a understand disability data because, that is the important aspect of this ProTherm database. So, you want pro thermodynamic data ok.

So, now the display option you just clear it clear of everything right. So, if you click on the delta delta G H 2 O all right, this is the change in the free energy obtained denaturant. So, if you do so you have to make sure that you should not click the thermal here, if you click the thermal and if you do that delta delta G H 2 O, you do not get any data right your denaturant that is fine, if you do the thermal, then you have to look for the delta Tm, or the delta delta G not with the H 2 O right the make sure ok.

If you do like this right, if you start like this what will happen you get this data.

(Refer Slide Time: 10:47)



Entry	ddG_H2O
2	-3.00
3	-0.60
4	-3.10
7	NULL
8	NULL
9	NULL
10	NULL
11	NULL
12	NULL
13	NULL
14	NULL
15	NULL
16	NULL
18	-2.10
19	-3.10
20	-2.80
21	-3.70
22	-3.10
23	-2.10
24	1.50
26	NULL
27	NULL
28	NULL
29	NULL

Now, the problem is you get the data, but you do not know this data corresponds to what right. So, this is not sufficient so, what are the important information necessary for the analysis, what information necessary.

Student: (Refer Time: 11:03).

You know protein m right, or if you are interested in the structures, then we need the PDB while, then it will tell you whether the mutant three structure, or not right, then you can make the energy structure is known, you can know structure information, structure information not known you can do sequence information.

Then whatever that information do you need.

Student: (Refer Time: 11:22).

Yeah this is very important right otherwise we do not know the mutation (Refer Time: 11:27) very important that will tell you, work which residue is mutated to which residue in which protein this is fine and anything else do you want.

Student: (Refer Time: 11:37).

Delta G if you talk about the a denaturant right, if you interested in denaturants denaturation delta G we do not get any data. So, you get the, a delta G H 2 O right.

So, delta G H 2 be had, then we can also think about your conditions right, because their data varies from different conditions. So, you can use the T temperature we get denaturant right ok, I can also click at the if you interested in denaturant right, then you can see the pH is these right ok. Let us see what is what is happening here.

(Refer Slide Time: 12:15)

Entry	Protein	PDB	wild Mutation	Sec.Str	ASA(%)dG	H ₂ O dHG	H ₂ O T	m	C _m	pH	
2	Phospholipase A2	1BP2	H 48 N	H	9.40	6.50	-3.00	30.00	1.20	5.40	8.00
3	Phospholipase A2	1BP2	H 48 Q	H	9.40	8.90	-0.60	30.00	1.34	6.60	8.00
4	Phospholipase A2	1BP2	H 48 A	H	9.40	6.40	-3.10	30.00	1.02	6.30	8.00
18	Interleukin 1 beta	1I0B	T 9 A	S	3.89	7.00	-2.10	25.00	NULL	1.10	6.50
19	Interleukin 1 beta	1I0B	T 9 L	S	3.89	6.00	-3.10	25.00	NULL	1.20	6.50
20	Interleukin 1 beta	1I0B	T 9 Q	S	3.89	6.30	-2.80	25.00	NULL	1.00	6.50
21	Interleukin 1 beta	1I0B	T 9 G	S	3.89	5.40	-3.70	25.00	NULL	0.90	6.50
22	Interleukin 1 beta	1I0B	K 97 R	H	44.48	6.00	-3.10	25.00	NULL	1.20	6.50
23	Interleukin 1 beta	1I0B	K 97 G	H	44.48	7.00	-2.10	25.00	NULL	1.10	6.50
24	Interleukin 1 beta	1I0B	K 97 V	H	44.48	10.60	1.50	25.00	NULL	1.40	6.50
26	Gene V	1X0B	E 30 F	S	16.94	NULL	NULL	25.00	3.60	3.18	7.00
27	Gene V	1X0B	E 30 M	S	16.94	NULL	NULL	25.00	3.70	2.78	7.00
28	Gene V	1X0B	C 33 S	S	0.00	NULL	NULL	25.00	4.00	1.40	7.00
29	Gene V	1X0B	D 36 C	S	34.42	NULL	NULL	25.00	3.50	2.00	7.00
30	Gene V	1X0B	D 36 N	S	34.42	NULL	NULL	25.00	4.20	2.20	7.00
31	Gene V	1X0B	I 47 T	S	2.43	NULL	NULL	25.00	4.90	0.90	7.00
32	Gene V	1X0B	D 50 H	T	28.52	NULL	NULL	25.00	3.80	2.10	7.00
33	Gene V	1X0B	F 68 L	S	31.39	NULL	NULL	25.00	3.40	1.40	7.00
34	Gene V	1X0B	K 69 H	S	42.15	NULL	NULL	25.00	3.70	2.20	7.00
35	Gene V	1X0B	K 69 M	S	42.15	NULL	NULL	25.00	3.10	2.70	7.00
36	Gene V	1X0B	V 70 C	T	68.05	NULL	NULL	25.00	4.40	1.70	7.00
37	Gene V	1X0B	V 70 P	T	68.05	NULL	NULL	25.00	4.40	1.30	7.00
38	Gene V	1X0B	V 70 T	T	34.05	8.01	-0.81	25.00	1.88	4.20	6.30

So, I got the data like you know this is a compatible because, it is not exiting the page limit and all right. So, you can see this the protein m this is a pretty video all most from all the mutants, they have the wild type IID, if you want to have the mutant ID they are click on mutant, then we know that success of the wild type and mutant.

Then you come to the compare two analysis right ok, this is secondary structure right. Now, these proteins if you want to you have any analysis based on secondary structures, you can extract right we discussed about AWK, just you can save it and you receive a single AWK come and, you can get the secondary information right.

Then in this case also if you see there are many null null data ok, either remove missing AWK comment. Otherwise in the database also you can eliminate these how to eliminate these right, we going to the search option go to delta energy H 2 O, we should get some numbers. So, you can give two extreme values server actually can give minus 100 to plus 100 ok.

So, m and c at the moment really we (Refer Time: 13:29) that. So, in this case if you see this you see delta delta G H 2 O you get the numbers right, you can eliminate all the null values. So, if you want to remove the null values you can give some range, then this case you got (Refer Time: 13:45) element because, it will show on only you figure the numbers right.

So, now you get the some of the data right now, if you the if you see this right some cases the same mutants appears several times, that will be difficult to search with this one. So, in this case you can use the sorting option ok. So, if you see this one the sorting option.

If you do with this wild type residue and the mutant residue and, the residue number ok, then you can also use the values because, delta G H 2 you are now using on so, delta G H 2, then you will see same mutation maybe same protein or different protein and different locations, how we will change right, you can see that either taken to the ascending order either taken to the descending order right, you may be descending order this is better because, we get the stabilizing ones the first right.

(Refer Slide Time: 14:45)

Entry	Protein	PDB	wild	Mutation	Sec.Str	ASA(%)	dG	H2O	dG	H2O T	pH
5352	Phosphoglycerate kinase	3PGK	Y	122 W	C	37.67	4.81	-0.57	25.00	7.50	
5357	Phosphoglycerate kinase	3PGK	Y	122 W	C	37.67	3.45	-0.91	25.00	7.50	
5365	Phosphoglycerate kinase	3PGK	Y	122 W	C	37.67	5.79	-1.29	25.00	7.50	
5374	Phosphoglycerate kinase	3PGK	Y	122 W	C	37.67	2.72	-5.65	25.00	7.50	
5351	Phosphoglycerate kinase	3PGK	Y	48 W	H	37.58	5.31	-0.07	25.00	7.50	
5356	Phosphoglycerate kinase	3PGK	Y	48 W	H	37.58	3.80	-0.56	25.00	7.50	
5364	Phosphoglycerate kinase	3PGK	Y	48 W	H	37.58	4.31	-1.77	25.00	7.50	
5371	Phosphoglycerate kinase	3PGK	Y	48 W	H	37.58	3.54	-4.83	25.00	7.50	
12082	CcpA	1JMC	Y	42 W	NULL	NULL	3.20	0.20	25.00	7.00	
12086	CcpA	1JMC	Y	42 W	NULL	NULL	3.20	0.10	25.00	7.00	
2658	Staphylococcal nuclease	1STN	Y	27 W	S	10.29	4.40	-0.60	20.00	7.00	
2657	Staphylococcal nuclease	1STN	Y	27 Y	S	10.29	2.00	-3.00	20.00	7.00	
8789	U1A protein	1OLA	Y	86 T	S	6.46	5.20	-2.90	22.00	7.00	
2656	Staphylococcal nuclease	1STN	Y	27 T	S	10.29	1.70	-3.30	20.00	7.00	
4593	Inulinin	1ZNF	Y	26 T	S	2.01	3.46	-0.30	25.00	8.00	
8733	U1A protein	1OLA	Y	13 T	S	19.00	6.80	-1.30	22.00	7.00	
12222	Ketosteroid isomerase	1QHO	Y	55 S (PDB: Y 57 S; PIR: Y 57 S)	H	5.71	14.80	9.50	25.00	7.00	
8735	U1A protein	1OLA	Y	31 S	H	43.14	5.80	-2.30	22.00	7.00	
12226	Ketosteroid isomerase	1QHO	Y	30 S (PDB: Y 32 S; PIR: Y 32 S)	S	0.00	10.60	13.70	25.00	7.00	
2655	Staphylococcal nuclease	1STN	Y	27 S	S	10.29	2.00	-3.00	20.00	7.00	
1678	Barnase	1BNT	Y	17 S	H	56.62	6.80	-2.00	25.00	6.30	
12235	Ketosteroid isomerase	1QHO	Y	14 S (PDB: Y 16 S; PIR: Y 16 S)	H	3.18	12.40	11.90	25.00	7.00	

So, we start its interesting you can see the delta delta G ok, this is the or because of this descending order we get this Y here right. So, in the, if you put the ascending order let us see level ok.

(Refer Slide Time: 15:01)

ProTherm search result

www.abren.net/protherm/protherm_result.php

Search Condition
Mutation No.:Single,
Method:Denaturants,
dG_H2O >:100 to 100
Sorting by wild_res.amut_res.res.no.ddg_h2o.

NEXT

Entry Protein	PDB_wild Mutation	Sec.Str	ASA(%)	dG_H2O	ddG_H2O T	pH
188 Barnase	JBNJ A 32 C	H	83.76	6.98	-1.85	25.00 6.30
9201 Staphylococcal nuclease	IJSTN A 60 C	H	55.08	4.30	-1.20	20.00 7.00
4611 Fatty acid binding protein	2JFB A 104 C	E	3.54	7.10	1.51	20.00 9.60
4604 Fatty acid binding protein	2JFB A 104 C	S	3.54	7.76	2.54	20.00 7.20
9207 Staphylococcal nuclease	IJSTN A 112 C	C	17.51	4.70	-0.80	20.00 7.00
8398 Coiled-coil protein	NULL A 22 D	NULL	NULL	1.23	-1.80	25.00 7.00
180 Barnase	JBNJ A 32 D	H	83.76	8.42	-0.41	25.00 6.30
8399 Coiled-coil protein	NULL A 22 E	NULL	NULL	0.33	-2.70	25.00 7.00
8381 Acylphosphatase	NULL A 28 E	NULL	NULL	4.45	0.41	28.00 5.50
127 Barnase	JBNJ A 32 E	H	83.76	8.96	0.13	25.00 6.30
74907 Staphylococcal nuclease	IJSTN A 58 E	H	0.00	5.00	-6.90	20.00 8.00
74908 Staphylococcal nuclease	IJSTN A 58 E	H	0.00	4.70	-6.90	20.00 9.00
74909 Staphylococcal nuclease	IJSTN A 58 E	H	0.00	3.90	-6.40	20.00 9.90
74906 Staphylococcal nuclease	IJSTN A 58 E	H	0.00	6.90	-4.80	20.00 5.90
74905 Staphylococcal nuclease	IJSTN A 58 E	H	0.00	7.50	-4.50	20.00 4.90
74904 Staphylococcal nuclease	IJSTN A 58 E	H	0.00	5.70	-3.80	20.00 3.90
74910 Staphylococcal nuclease	IJSTN A 90 E	C	0.00	3.30	-8.60	20.00 8.00
74911 Staphylococcal nuclease	IJSTN A 90 E	C	0.00	3.30	-8.20	20.00 9.10
74920 Staphylococcal nuclease	IJSTN A 90 E	C	0.00	4.00	-7.70	20.00 5.90
74921 Staphylococcal nuclease	IJSTN A 90 E	C	0.00	1.90	-7.50	20.00 3.90
74922 Staphylococcal nuclease	IJSTN A 90 E	C	0.00	4.70	-7.10	20.00 4.90
74967 Staphylococcal nuclease	IJSTN A 109 E	C	0.00	3.40	-8.10	20.00 9.10

So, here if you see the same mutations A to C, because alanine is the first one then you put C right. So, say different protein is barnase and they sub nucleus fatty acid binding protein, the here is these two the same data, the these two A 1 C 104 C 104 C same protein maybe different conditions.

This way secondary structure right, you can see this a values are different is 1.251 1 point we get pH is different is 9.6 this 1.2 right.

(Refer Slide Time: 15:34)

ProTherm search result

www.abren.net/protherm/protherm_result.php

74909 Staphylococcal nuclease	IJSTN A 58 E	H	0.00	3.90	-6.40	20.00 9.90
74906 Staphylococcal nuclease	IJSTN A 58 E	H	0.00	6.90	-4.80	20.00 5.90
74905 Staphylococcal nuclease	IJSTN A 58 E	H	0.00	7.50	-4.50	20.00 4.90
74904 Staphylococcal nuclease	IJSTN A 58 E	H	0.00	5.70	-3.80	20.00 3.90
74910 Staphylococcal nuclease	IJSTN A 90 E	C	0.00	3.30	-8.60	20.00 8.00
74911 Staphylococcal nuclease	IJSTN A 90 E	C	0.00	3.30	-8.20	20.00 9.10
74920 Staphylococcal nuclease	IJSTN A 90 E	C	0.00	4.00	-7.70	20.00 5.90
74921 Staphylococcal nuclease	IJSTN A 90 E	C	0.00	1.90	-7.50	20.00 3.90
74922 Staphylococcal nuclease	IJSTN A 90 E	C	0.00	4.70	-7.10	20.00 4.90
74967 Staphylococcal nuclease	IJSTN A 109 E	C	0.00	3.40	-8.10	20.00 9.10
74966 Staphylococcal nuclease	IJSTN A 109 E	C	0.00	4.20	-7.70	20.00 7.90
74965 Staphylococcal nuclease	IJSTN A 109 E	C	0.00	6.20	-5.50	20.00 6.00
74964 Staphylococcal nuclease	IJSTN A 109 E	C	0.00	6.80	-5.00	20.00 5.00
74963 Staphylococcal nuclease	IJSTN A 109 E	C	0.00	5.10	-4.40	20.00 3.90
74978 Staphylococcal nuclease	IJSTN A 132 E	H	0.00	3.00	-8.50	20.00 9.00
74977 Staphylococcal nuclease	IJSTN A 132 E	H	0.00	3.70	-8.20	20.00 7.90
74976 Staphylococcal nuclease	IJSTN A 132 E	H	0.00	4.90	-6.80	20.00 6.00
74975 Staphylococcal nuclease	IJSTN A 132 E	H	0.00	6.00	-5.80	20.00 4.90
8388 Coiled-coil protein	NULL A 22 F	NULL	NULL	4.23	1.20	25.00 7.00
179 Barnase	JBNJ A 32 F	H	83.76	8.40	-0.43	25.00 6.30
3131 Staphylococcal nuclease	IJSTN A 60 F	H	55.08	4.70	-0.80	20.00 7.00
3148 Staphylococcal nuclease	IJSTN A 102 F	H	16.42	3.60	-1.90	20.00 7.00
3150 Staphylococcal nuclease	IJSTN A 112 F	C	17.51	3.90	-1.60	20.00 7.00
9215 Staphylococcal nuclease	IJSTN A 112 F	C	17.51	4.20	-1.30	20.00 7.00
11591 Alpha1-antitrypsin	IQLP A 183 F	NULL	NULL	NULL	1.50	25.00 6.50
11595 Alpha1-antitrypsin	IQLP A 248 F	NULL	NULL	NULL	1.80	25.00 6.50
6985 Acyl-coenzyme a binding protei	2ABD A 9 G	H	8.17	6.08	-2.00	5.00 5.30
1988 Staphylococcal nuclease	IJSTN A 12 G	C	13.52	3.10	-2.40	20.00 7.00
8435 Fibronectin	1TGU A 13 G	S	30.22	NULL	-0.93	25.00 5.00

All also you can see this A 132 E so, you can see the difference in values minus 8.5 to minus 5.8 ok. This is major difference it is dependent upon the different pH the effect of pH.

So, if you have the delta g values, a same mutation from different pH, then you can do the analysis what is the effect of pH on it is particular mutation. If you can generalize for the different proteins, then we can make an analyses ok, this is the case for the general generally for different proteins, you can do analyses and to see whether you can generalize, or we cannot generalize depending upon the protein or not right.

(Refer Slide Time: 16:14)

Accession	Protein Name	Mutation	Residue	Type	pH 5.8	pH 7.0	pH 8.5	pH 10.0
9236	Barnase	JBSI	D 8 A	H	61.00	8.94	-0.99	25.00 6.30
44	Barnase	JBSI	D 8 A	H	61.00	7.97	-0.85	25.00 6.30
18684	Alkaline proteinase inhibitor	NULL	D 10 A	H	45.87	6.50	2.70	25.00 7.10
7017	Ribonuclease HI	2RNS	D 10 A	S	12.77	13.00	3.30	25.00 5.50
47	Barnase	JBSI	D 12 A	H	54.20	8.43	-0.39	25.00 6.30
2472	Barnase	JBSI	D 12 A	H	54.20	NULL	-0.34	25.00 6.30
9237	Barnase	JBSI	D 12 A	H	54.20	9.59	-0.34	25.00 6.30
3683	Barnase	JBSI	D 12 A	H	54.20	10.12	0.11	25.00 6.30
18985	Ribosomal protein L30E	JITM	D 12 A	H	80.50	11.60	0.24	25.00 7.40
9787	Chemotactic protein	JCFV	D 12 A	C	3.19	8.10	2.50	25.00 7.00
9788	Chemotactic protein	JCFV	D 13 A	C	40.04	8.30	2.70	25.00 7.00
24398	Aspartate aminotransferase	JAMQ	D 15 A	NULL	47.08	NULL	-5.90	25.00 7.50
24397	Aspartate aminotransferase	JAMQ	D 15 A	NULL	47.08	NULL	-5.80	25.00 7.50
24399	Aspartate aminotransferase	JAMQ	D 15 A	NULL	47.08	NULL	-5.10	25.00 7.50
18143	Spectrin	JAMJ	D 16 A (PDB: D 17 A; PIR: D 1779 A)	H	28.24	6.10	-0.20	25.00 7.00
1565	Staphylococcal nuclease	JSTN	D 19 A	T	3.40	5.50	-0.10	20.00 7.00
8685	Staphylococcal nuclease	JSTN	D 19 A	T	6.80	5.30	-0.10	20.00 7.00
92	Arc repressor	JARR	D 20 A	H	36.22	10.10	-0.80	25.00 7.50
6089	Acyl-coenzyme A binding protein	JABD	D 21 A	H	18.46	7.68	-0.40	5.00 5.30
1567	Staphylococcal nuclease	JSTN	D 21 A	T	6.80	6.30	0.70	20.00 7.00
8683	Staphylococcal nuclease	JSTN	D 21 A	T	6.80	6.10	0.70	20.00 7.00
16445	Ribosomal protein L9	JDTV	D 23 A	T	67.87	4.30	-0.51	25.00 5.40
16441	Ribosomal protein L9	JDTV	D 23 A	T	67.87	3.99	-0.46	25.00 5.40
16446	Ribosomal protein L9	JDTV	D 23 A	T	67.87	3.53	1.13	25.00 5.40
16442	Ribosomal protein L9	JDTV	D 23 A	T	67.87	3.27	1.25	25.00 5.40
4001	Thioredoxin	JTRX	D 26 A	S	0.76	13.20	3.70	25.00 7.00

So, now, if you see this you can see the display is more than three 100 right in this case if you want to change it right here, you can increase the number and, if you start it takes a time because we need to not everything it is not. So, slow so you can get the unity it is more because salaries, it is displayed 1 to 3000. So, you have more number of data and you can increase a numbers and do that ok.

So, now, the in this case if you are interested more details on the first mutation barnase here, we will going to any information. So, what you do in this case the display option, you also click on the reference right and, if you are interested in the mutant structures, then we take click on mutant also right, then if you going to starts right then reduce this number.

(Refer Slide Time: 17:05)

ProTherm search result

www.abren.net/protherm/protherm_result.php

Search Condition
Mutation No.:Single,
Method:Denaturants,
dG_H2O >:100 to 100
Sorting by wild_res.amnt_res.res.no.ddg_h2o.

NEXT

Entry Protein	PDB_wild	PDB_mutant	Mutation	Sec.Str	ASA(%)	dG_H2O	ddG_H2O	T	pH	REFERENCE
188 Barnase	1BNI	NULL	A 32 C	H	83.76	6.98	-1.85	25.00	6.30	J.MOL.BIOL. 227, 560-568 (1992) PMID: 1404569
9203 Staphylococcal nuclease	1STN	NULL	A 60 C	H	55.08	4.30	-1.20	20.00	7.00	PROTEIN SCI 4, 2545-2558 (1995) PMID: 5580845
4611 Fatty acid binding protein	2IFB	NULL	A 104 C	E	3.54	7.10	1.51	20.00	9.60	BIOCHEMISTRY 37, 11015-11021 (1993) PMID: 8218166
4604 Fatty acid binding protein	2IFB	NULL	A 104 C	S	3.54	7.76	2.54	20.00	7.20	BIOCHEMISTRY 37, 11015-11021 (1993) PMID: 8218166
9207 Staphylococcal nuclease	1STN	NULL	A 112 C	C	17.51	4.70	-0.80	20.00	7.00	PROTEIN SCI 4, 2545-2558 (1995) PMID: 5580845
8399 Coiled-coil protein	NULL	NULL	A 22 D	NULL	NULL	1.23	-1.80	25.00	7.00	J.MOL.BIOL. 300, 377-402 (2000) PMID: 10874472
180 Barnase	1BNI	NULL	A 32 D	H	83.76	8.42	-0.41	25.00	6.30	J.MOL.BIOL. 227, 560-568 (1992) PMID: 1404569
8399 Coiled-coil protein	NULL	NULL	A 22 E	NULL	NULL	0.33	-2.70	25.00	7.00	J.MOL.BIOL. 300, 377-402 (2000) PMID: 10874472
8381 Acylphosphatase	NULL	NULL	A 28 E	NULL	NULL	4.45	0.41	28.00	5.50	J.MOL.BIOL. 300, 633-647 (2000) PMID: 10884358
177 Barnase	1BNI	NULL	A 32 E	H	83.76	8.96	0.13	25.00	6.30	J.MOL.BIOL. 227, 560-568 (1992) PMID: 1404569
24907 Staphylococcal nuclease	1STN	NULL	A 58 E	H	0.00	5.00	-6.90	20.00	8.00	PROC.NATL.ACAD.SCI.U.S.A 107, 16096-16100 (2010) PMID: 20783341
24908 Staphylococcal nuclease	1STN	NULL	A 58 E	H	0.00	4.70	-6.90	20.00	9.00	PROC.NATL.ACAD.SCI.U.S.A 107, 16096-16100 (2010) PMID: 20783341

Now, this number is not necessary right. So, you reduce 300 so, we will see these the reference right and the some of the cases PD mutant is most (Refer Time: 17:17) of element if it is lysozyme you can get the a mutant a data.

So, I cannot is it is nothing is not to displayed here you can see the mutant structure right.

(Refer Slide Time: 17:28)

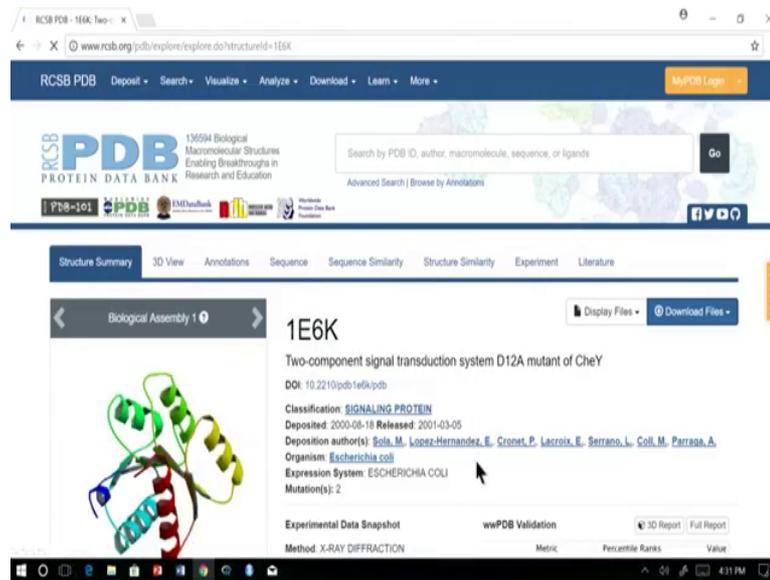
ProTherm search result

www.abren.net/protherm/protherm_result.php

Entry Protein	PDB_wild	PDB_mutant	Mutation	Sec.Str	ASA(%)	dG_H2O	ddG_H2O	T	pH	REFERENCE
47 Barnase	1BNI	NULL	D 12 A	H	54.20	8.43	-0.39	25.00	6.30	J.MOL.BIOL. 224, 783-804 (1992) PMID: 1569557
2472 Barnase	1BNI	NULL	D 12 A	H	54.20	NULL	-0.34	25.00	6.30	J.MOL.BIOL. 224, 733-740 (1992) PMID: 1569552
9232 Barnase	1BNI	NULL	D 12 A	H	54.20	9.59	-0.34	25.00	6.30	J.MOL.BIOL. 216, 1031-1044 (1990) PMID: 2266554
3683 Barnase	1BNI	NULL	D 12 A	H	54.20	10.12	0.11	25.00	6.30	BIOCHEMISTRY 29, 9343-9352 (1990) PMID: 2248951
18985 Ribosomal protein L30E	1H7M	NULL	D 12 A	H	80.50	11.60	0.24	25.00	7.40	BIOCHEMISTRY 44, 16817-16825 (2005) PMID: 16363795
9787 Chemotactic protein	1CEY	1E6K	D 12 A	C	3.19	8.10	2.50	25.00	7.00	J.MOL.BIOL. 303, 213-225 (2000) PMID: 11023787
9788 Chemotactic protein	1CEY	1E6L	D 13 A	C	40.04	8.30	2.70	25.00	7.00	J.MOL.BIOL. 303, 213-225 (2000) PMID: 11023787
24398 Aspartate aminotransferase	1AMQ	NULL	D 15 A	NULL	47.08	NULL	-5.90	25.00	7.50	BIOCHEMISTRY 48, 433-441 (2009) PMID: 19099421
24397 Aspartate aminotransferase	1AMQ	NULL	D 15 A	NULL	47.08	NULL	-5.80	25.00	7.50	BIOCHEMISTRY 48, 433-441 (2009) PMID: 19099421
24399 Aspartate aminotransferase	1AMQ	NULL	D 15 A	NULL	47.08	NULL	-5.10	25.00	7.50	BIOCHEMISTRY 48, 433-441 (2009) PMID: 19099421
18142 Spectrin	1A13	NULL	D 16 A (PDB: D 17 A; PIR: D 179 A)	H	28.24	6.10	-0.20	25.00	7.00	J.MOL.BIOL. 344, 207-221 (2004) PMID: 15704117
1565 Staphylococcal nuclease	1STN	NULL	D 19 A	T	3.40	5.50	-0.10	20.00	7.00	BIOCHEMISTRY 35, 6443-6449 (1996) PMID: 8639591
8685 Staphylococcal nuclease	1STN	NULL	D 19 A	T	6.80	5.30	-0.10	20.00	7.00	J.MOL.BIOL. 235, 27-37 (1994) PMID: 8289248
592 Arc repressor	1ARB	NULL	D 20 A	H	36.22	10.10	-0.80	25.00	7.50	NAT.STRUCT.BIOL. 1, 518-523 (1994) PMID: 7664079
6999 Acyl-coenzyme A binding protein	2ABD	NULL	D 21 A	H	18.46	7.68	-0.40	5.00	5.30	NAT.STRUCT.BIOL. 6, 594-601 (1999) PMID: 10760367
			D 21 A	T	6.80	6.30	0.70	20.00	7.00	BIOCHEMISTRY 35, 6443-6449 (1996) PMID: 8639591

If you click on this one E 6 L here a 1 E 6 k it is a pretty body.

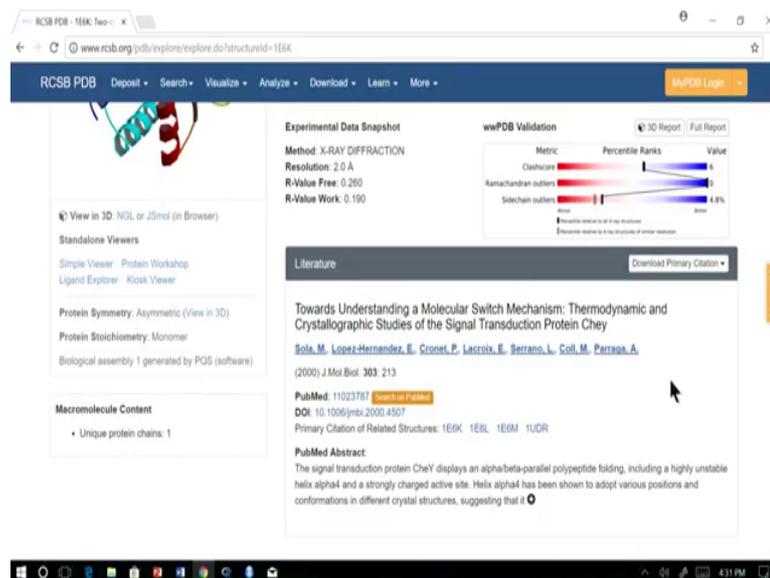
(Refer Slide Time: 17:36)



The screenshot shows the RCSB PDB website interface. The main header includes navigation links: Deposit, Search, Visualize, Analyze, Download, Learn, and More. A search bar is present with the text "Search by PDB ID, author, macromolecule, sequence, or ligands". Below the header, the PDB logo is displayed with the text "136594 Biological Macromolecular Structures Enabling Breakthroughs in Research and Education". The main content area shows the entry for 1E6K, titled "Two-component signal transduction system D12A mutant of CheY". The entry includes a 3D ribbon diagram of the protein structure. Key information provided includes: DOI: 10.2210/pdb/1e6k/pdb, Classification: SIGNALING PROTEIN, Deposited: 2000-08-18, Released: 2001-03-05, Deposition author(s): [Sola, M.](#), [Lopez-Hernandez, E.](#), [Cronet, P.](#), [Lacroix, E.](#), [Serrano, L.](#), [Coll, M.](#), [Parraga, A.](#), Organism: [Escherichia coli](#), Expression System: ESCHERICHIA COLI, Mutation(s): 2. The experimental method is listed as X-RAY DIFFRACTION. A wwPDB Validation section is also visible, showing a table with columns for Metric, Percentile Ranks, and Value.

So, you can see this is the mutant all right.

(Refer Slide Time: 17:41)



This screenshot shows a more detailed view of the 1E6K entry on the RCSB PDB website. The left sidebar contains options for viewing the structure in 3D (NGL or JSmol) and standalone viewers (Simple Viewer, Protein Workshop, Ligand Explorer, Kiosk Viewer). It also lists protein symmetry (Asymmetric), stoichiometry (Monomer), and biological assembly (1 generated by PQS). The main content area features an "Experimental Data Snapshot" table with the following data: Method: X-RAY DIFFRACTION, Resolution: 2.0 Å, R-Value Free: 0.260, and R-Value Work: 0.190. To the right, a "wwPDB Validation" section displays a bar chart comparing metrics: Clashscore (Value: 6), Ramachandran outliers (Value: 4.8%), and Sidechain outliers (Value: 4.8%). Below this, the "Literature" section shows a primary citation: "Towards Understanding a Molecular Switch Mechanism: Thermodynamic and Crystallographic Studies of the Signal Transduction Protein CheY" by [Sola, M.](#), [Lopez-Hernandez, E.](#), [Cronet, P.](#), [Lacroix, E.](#), [Serrano, L.](#), [Coll, M.](#), [Parraga, A.](#) (2000) *J Mol Biol* 303: 213. The PubMed ID is 11023787, and the DOI is 10.1006/jmbi.2000.4507. A PubMed abstract is also visible, describing the protein's structure and function.

So, you can see this right. So, you can see the, this is the wild type and this is the mutation and, you can see this one.

(Refer Slide Time: 17:50)

Accession	Protein	Substitution	Position	Class	ΔG (kcal/mol)	ΔTm (°C)	ΔTm (°C)	ΔTm (°C)	ΔTm (°C)	Reference	
3048	Tumor suppressor P53 complexed	L1UP	NULL	C 242 S	C	13.60	6.82	-2.94	10.00	7.20	PROC NATL ACAD SCI U S A 94 14336 14342 (1997) PMID: 9405611
4786	Cytochrome c	LYCC	NULL	C 102 T	C	0.85	6.40	2.30	27.00	7.20	BIOCHEMISTRY 31 12337-12344 (1992) PMID 1334426
7167	Iso-1 cytochrome c	LYCC	NULL	C 102 T	C	0.85	6.48	4.09	30.00	7.00	EUR J BIOCHEM 223 155-160 (1994) PMID: 8033883
4325	Lysozyme	LIZM	NULL	C 54 Y	T	2.64	NULL	-5.02	22.00	7.00	BIOCHEMISTRY 30 589-594 (1991) PMID: 1985046
18951	Ribosomal protein L30e	LHTM	NULL	D 2 A	C	56.63	10.84	-0.53	25.00	7.40	BIOCHEMISTRY 44 16817-16825 (2005) PMID 16363795
7471	Barnase	LBSN	NULL	D 8 A	H	61.00	NULL	-0.99	25.00	6.30	J MOL BIOL 224 733-740 (1992) PMID: 1569552
9236	Barnase	LBSN	NULL	D 8 A	H	61.00	8.94	-0.99	25.00	6.30	J MOL BIOL 216 1031-1044 (1990) PMID: 2266554
44	Barnase	LBSN	NULL	D 8 A	H	61.00	7.97	-0.85	25.00	6.30	J MOL BIOL 224 783-804 (1992) PMID: 1569552
18684	Alkaline proteinase inhibitor	NULL	NULL	D 10 A	H	45.87	6.50	2.70	25.00	7.10	BIOCHEMISTRY 44 2469-2477 (2005) PMID: 15309759
7017	Ribonuclease HI	LBSN	NULL	D 10 A	S	12.77	13.00	3.30	25.00	5.50	NAT STRUCT BIOL 6 825-831 (1999) PMID: 10467093
47	Barnase	LBSN	NULL	D 12 A	H	54.20	8.43	-0.39	25.00	6.30	J MOL BIOL 224 783-804 (1992) PMID: 1569552
7472	Barnase	LBSN	NULL	D 12 A	H	54.20	NULL	-0.34	25.00	6.30	J MOL BIOL 224 733-740 (1992) PMID: 1569552
9237	Barnase	LBSN	NULL	D 12 A	H	54.20	9.59	-0.34	25.00	6.30	J MOL BIOL 216 1031-1044 (1990) PMID: 2266554
3683	Barnase	LBSN	NULL	D 12 A	H	54.20	10.12	0.11	25.00	6.30	BIOCHEMISTRY 29 9343-9352 (1990) PMID: 2248951
18952	Ribosomal protein L30e	LHTM	NULL	D 12 A	H	80.50	11.60	0.24	25.00	7.40	BIOCHEMISTRY 44 16817-16825 (2005) PMID 16363795
				2 A	C	3.19	8.10	2.50	25.00	7.00	J MOL BIOL 303 213-225 (2000) PMID: 11233727

So, if you want to get the complete information you will go to the purple right. So, in the bio chemistry paper, if you see the purple you can access the abstract.

(Refer Slide Time: 18:03)

Effects of charge-to-alanine substitutions on the stability of ribosomal protein L30e from *Thermococcus celer*.

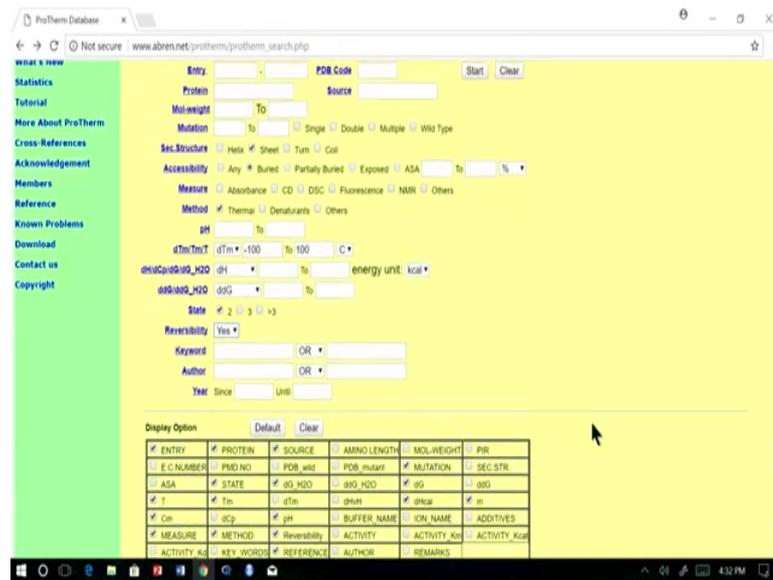
Lee CF¹, Mahtabatzke GJ, Wong KB

Abstract
The ability to rationally engineer a protein with altered stability depends upon the detailed understanding of the role of noncovalent interactions in defining thermodynamic properties of proteins. In this paper, we used *T. celer* L30e as a model to address the question of the role of charge-charge interactions in defining the stability of this protein. A total of 26 single-site charge-to-alanine variants of this protein were generated, and the stability of these proteins was determined using thermal- and denaturant-induced unfolding. It was found that, although L30e is isolated from a thermophilic organism and is highly thermostable, some of the substitutions lead to a further increase in the transition temperature. Analysis of the effects of high ionic strength on the stabilities of L30e variants shows that the long-range charge-charge interactions are as important as the short-range (salt bridge) interactions. The changes in stabilities of the *T. celer* L30e protein variants were compared with the changes in the energy of charge-charge interactions calculated using different computational models. It was found that there is a good qualitative agreement between experimental and calculated data: for 70-80% (19-21 of 26, confidence $p < 0.003$) of the variants, computational models predict correctly the sign of the stability changes. In particular, computational models identify correctly those charged amino acid residue substitutions of which led to enhancement in thermostability. Thus, optimization of the charge-charge interactions might be a useful approach for the rational increase in protein stability.

PMID 16363795 DOI 10.1021/bk-05-0864
(Indexed for MEDLINE)

And if you have access then you can see the full text right, then you can get all this information right fine right.

(Refer Slide Time: 18:11)



So, now if you want have multiple a choices for example, you are interested only on the buried mutations and in sheets and, obtained with the a thermal denaturation right, then you better to select the dTm, there is a range is minus 1 100 to plus 100 degree C right.

You can see centigrade or Kelvin here also, you can is kilo kcal kilo jule so, (Refer Time: 18:45) so, what is the conversion between kilo kcal and kilo jule right, when kilo kcal equal to 4.18 kilo jule right. So, we can see the number so, if you do like this right, we are simply to dH take S ok, this is commonly used one state 2 2 state values right.

Then if you do like this and here we need to change because, we already it is a thermal it automatically it a put the Tm and the T is not Ta C Tm, we do not need and measure method, we reduce it reversibility is yes is the reference fine, there is m we do not need delta delta G, we need and this we remove it and, if use source we take start oh right.

(Refer Slide Time: 19:32)

Search Condition
Sec. Str.:Sheet,
Accessibility:1-Buried
Method:Thermal,
dTm: -100 to 100State :2,
Reversibility: Yes

Entry	Protein	PDB_wild Mutation	State	dG	H2O dG	ddG	T	Tm	pH	REFERENCE
109	Staphylococcal nuclease	L25 A	2	NULL	NULL	NULL	NULL	39.90	7.00	PROTEINS SCL 1, 467-576 (1993) PMID: 8518370
115	Staphylococcal nuclease	L25 A	2	NULL	NULL	NULL	NULL	31.20	5.00	PROTEINS SCL 1, 467-576 (1993) PMID: 8518370
144	Ribonuclease HI	D10 N	2	NULL	NULL	NULL	NULL	53.80	9.00	BIOL CHEM 271, 3279-3276 1996) PMID: 8955106
145	Ribonuclease HI	D10 A	2	NULL	NULL	NULL	NULL	60.70	9.00	BIOL CHEM 271, 3279-3276 1996) PMID: 8955106
146	Ribonuclease HI	D10 E	2	NULL	NULL	NULL	NULL	50.40	9.00	BIOL CHEM 271, 3279-3276 1996) PMID: 8955106
147	Ribonuclease HI	D10 S	2	NULL	NULL	NULL	NULL	56.20	9.00	BIOL CHEM 271, 3279-3276 1996) PMID: 8955106
148	Ribonuclease HI	D10 H	2	NULL	NULL	NULL	NULL	55.20	9.00	BIOL CHEM 271, 3279-3276 1996) PMID: 8955106
147	Ribonuclease HI	D10 N	2	NULL	NULL	NULL	NULL	47.40	3.00	BIOL CHEM 271, 3279-3276 1996) PMID: 8955106
158	Ribonuclease HI	D10 A	2	NULL	NULL	NULL	NULL	58.10	3.00	BIOL CHEM 271, 3279-3276 1996) PMID: 8955106
159	Ribonuclease HI	D10 E	2	NULL	NULL	NULL	NULL	53.80	3.00	BIOL CHEM 271, 3279-3276 1996) PMID: 8955106
160	Ribonuclease HI	D10 S	2	NULL	NULL	NULL	NULL	52.40	3.00	BIOL CHEM 271, 3279-3276 1996) PMID: 8955106

So, delta yes delta delta G did I use this yes is thermal.

(Refer Slide Time: 19:49)

Known Problems
Download
Contact us
Copyright

Search filters:
pH: [] To []
dTm/Tm: [] To []
dTm: [] To []
dHcp/dHcp_H2O: [] To [] energy unit: kcal
ddG/dG_H2O: [] To []
State: [2] [3] [x3]
Reversibility: Yes *
Keyword: [] OR []
Author: [] OR []
Year: Since [] Until []

Display Option: Default Clear

<input checked="" type="checkbox"/> ENTRY	<input checked="" type="checkbox"/> PROTEIN	<input type="checkbox"/> SOURCE	<input type="checkbox"/> AMINO LENGTH	<input type="checkbox"/> MOLE WEIGHT	<input type="checkbox"/> PIR
<input type="checkbox"/> E.C NUMBER	<input type="checkbox"/> PMID NO	<input checked="" type="checkbox"/> PDB_wild	<input type="checkbox"/> PDB_mutant	<input type="checkbox"/> MUTATION	<input type="checkbox"/> SEC STR.
<input type="checkbox"/> ASA	<input checked="" type="checkbox"/> STATE	<input checked="" type="checkbox"/> dG_H2O	<input type="checkbox"/> dG_H2O	<input checked="" type="checkbox"/> dG	<input checked="" type="checkbox"/> dG
<input type="checkbox"/> T	<input type="checkbox"/> Tm	<input type="checkbox"/> dTm	<input type="checkbox"/> dHcp	<input type="checkbox"/> dHcp	<input type="checkbox"/> m
<input type="checkbox"/> Cn	<input type="checkbox"/> dCp	<input checked="" type="checkbox"/> pH	<input type="checkbox"/> BUFFER NAME	<input type="checkbox"/> ION NAME	<input type="checkbox"/> ADDITIVES
<input type="checkbox"/> MEASURE	<input type="checkbox"/> METHOD	<input type="checkbox"/> Reversibility	<input type="checkbox"/> ACTIVITY	<input type="checkbox"/> ACTIVITY Km	<input type="checkbox"/> ACTIVITY Kcat
<input type="checkbox"/> ACTIVITY kd	<input type="checkbox"/> KEY WORDS	<input type="checkbox"/> REFERENCE	<input type="checkbox"/> AUTHOR	<input type="checkbox"/> REMARKS	

Sorting By: OFF OFF OFF OFF ASCENDING

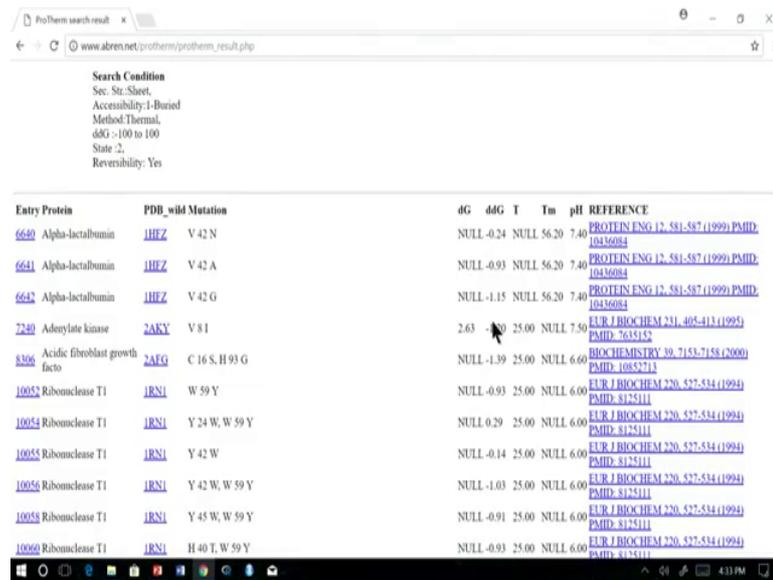
Entries per page: 300

Start Clear

Home | ProTherm | ProNIT
Biomolecules Gallery

Because I need the delta delta G this I now need, yes delta delta G you get the delta delta G values right for the thermal denaturants.

(Refer Slide Time: 20:16)



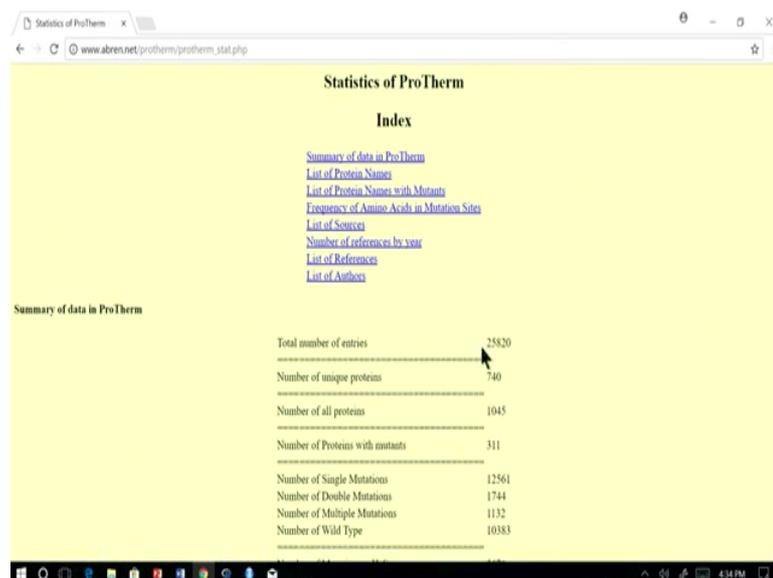
Search Condition
Sec. Str. Sheet,
Accessibility: 1-Buried
Method: Thermal,
dG > 100 to 100
State 2,
Reversibility: Yes

Entry	Protein	PDB	wild Mutation	dG	dG T	Tm	pH	REFERENCE
6640	Alpha-lactalbumin	1HFZ	V 42 N	NULL	-0.24	NULL	56.20	7.40 PROTEIN ENG 12, 581-587 (1999) PMID: 10436084
6641	Alpha-lactalbumin	1HFZ	V 42 A	NULL	-0.93	NULL	56.20	7.40 PROTEIN ENG 12, 581-587 (1999) PMID: 10436084
6642	Alpha-lactalbumin	1HFZ	V 42 G	NULL	-1.15	NULL	56.20	7.40 PROTEIN ENG 12, 581-587 (1999) PMID: 10436084
7240	Adenylate kinase	2AKY	V 81	2.63	-8.00	25.00	NULL	7.50 EUR J BIOCHEM 231, 405-413 (1995) PMID: 7635153
8306	Acidic fibroblast growth facto	2AFG	C 16 S, H 93 G	NULL	-1.39	25.00	NULL	6.60 BIOCHEMISTRY 39, 7153-7158 (2000) PMID: 10852713
10052	Ribonuclease T1	1RNI	W 59 Y	NULL	-0.93	25.00	NULL	6.00 EUR J BIOCHEM 220, 527-534 (1994) PMID: 8125111
10054	Ribonuclease T1	1RNI	Y 24 W, W 59 Y	NULL	0.29	25.00	NULL	6.00 EUR J BIOCHEM 220, 527-534 (1994) PMID: 8125111
10055	Ribonuclease T1	1RNI	Y 42 W	NULL	-0.14	25.00	NULL	6.00 EUR J BIOCHEM 220, 527-534 (1994) PMID: 8125111
10056	Ribonuclease T1	1RNI	Y 42 W, W 59 Y	NULL	-1.03	25.00	NULL	6.00 EUR J BIOCHEM 220, 527-534 (1994) PMID: 8125111
10058	Ribonuclease T1	1RNI	Y 45 W, W 59 Y	NULL	-0.91	25.00	NULL	6.00 EUR J BIOCHEM 220, 527-534 (1994) PMID: 8125111
10060	Ribonuclease T1	1RNI	H 40 T, W 59 Y	NULL	-0.93	25.00	NULL	6.00 EUR J BIOCHEM 220, 527-534 (1994) PMID: 8125111

So, you can use this information for the further or analysis as well as for the prediction right.

So, likewise you can the ProTherm you can use to end it up because, it contents if you take the statistics.

(Refer Slide Time: 20:31)



Statistics of ProTherm

Index

- [Summary of data in ProTherm](#)
- [List of Protein Names](#)
- [List of Protein Names with Mutants](#)
- [Frequency of Amino Acids in Mutation Sites](#)
- [List of Sources](#)
- [Number of references by year](#)
- [List of References](#)
- [List of Authors](#)

Summary of data in ProTherm

Total number of entries	25820
Number of unique proteins	740
Number of all proteins	1045
Number of Proteins with mutants	311
Number of Single Mutations	12561
Number of Double Mutations	1744
Number of Multiple Mutations	1132
Number of Wild Type	10383

So, it contents about 25000 entries that unfortunately current there is no current updates available ProTherm, but even then the available number of data are sufficient for the

analyses. Let us say it is 740 unique proteins right then, they would 311 proteins content of the mutants.

(Refer Slide Time: 20:46)

Statistics of ProTherm

www.abren.net/protherm/protherm_stat.php

Number of entries by Absorbance	1197
Number of entries of CD	11274
Number of entries of DSC	6416
Number of entries of Fluorescence	5650
Number of entries of NMR	985
Number of entries of Others	531

Number of entries of Thermal	15839
Number of entries of GdnHCl	5843
Number of entries of Urea	3931
Number of entries of Others	212

Total Number of References	1902

List of Protein Names

- 2-oxoglutarate dehydrogenase
- 3-chymotrypsin-like SARS-Cov m
- 3-isopropylmalate dehydrogenase
- 4-aminobutyrate aminotransferase
- 4-oxalotransferrin lactonase
- 5-enolpyruvyl shikimate-3-pho
- 7S globulin
- Acetylcholinesterase
- Abi protein
- Abrin II
- Acetylcholinesterase
- Acid phosphatase
- Acid phosphatase A
- Acidic calcium-binding protein
- Acidic fibroblast growth factor
- Actin
- Acyl-coenzyme A binding protein

And these are the statistics for the proteins and if you design, this is a protein names, you can list a protein names and the source names and so, on. So, within if you go to the frequency right.

(Refer Slide Time: 21:01)

Statistics of ProTherm

Frequency of Amino Acid Residues in Mutation Sites

www.abren.net/protherm/mutation.html

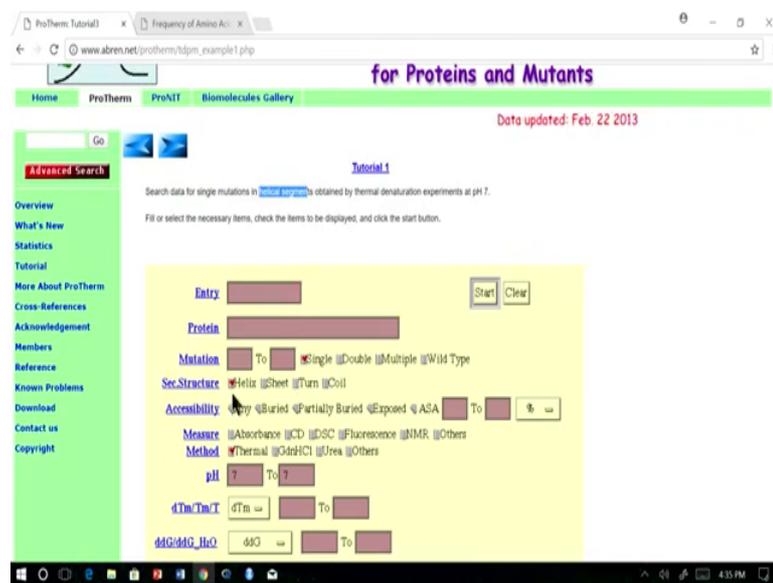
	To																			
	Gly	Ala	Val	Leu	Ile	Cys	Met	Phe	Trp	Pro	Ser	Thr	Asn	Gln	Asp	Glu	Lys	Arg	His	
Gly	---	230	57	5	0	12	0	10	2	6	14	53	2	5	26	28	24	6	19	13
Ala	113	---	131	61	20	29	24	21	5	3	85	74	44	10	18	14	13	41	5	11
Val	33	474	---	258	288	47	64	399	51	24	6	46	39	28	0	20	14	13	23	12
Leu	34	350	114	---	85	55	40	86	2	4	18	16	14	4	4	13	32	12	23	5
Ile	41	244	480	236	---	24	62	93	12	8	7	22	58	3	1	9	20	10	9	12
Cys	4	37	17	26	1	---	2	2	2	1	105	86	0	0	0	4	0	0	0	0
Met	18	63	50	124	45	1	---	17	2	0	0	0	6	1	0	6	6	18	21	0
Phe	5	149	24	117	18	3	10	---	95	63	0	24	5	21	2	4	4	7	0	7
Trp	31	39	4	25	2	20	1	185	---	72	4	11	4	6	8	17	1	4	2	9
Pro	0	31	1	18	0	3	133	69	---	0	2	0	1	5	3	3	0	3	13	3
Ser	66	187	7	17	2	9	0	4	2	4	---	68	10	4	4	6	5	5	8	3
Thr	28	212	19	18	10	29	2	13	6	2	8	---	28	20	4	66	6	30	37	25
Asn	37	204	126	32	103	38	11	24	28	2	3	118	---	20	27	18	92	7	40	21
Gln	34	153	6	6	51	6	12	4	0	1	1	33	10	---	5	97	22	13	11	41
Asp	42	74	3	23	5	12	3	3	4	0	9	5	1	14	---	10	34	33	14	6
Glu	58	199	8	9	9	33	4	14	7	8	12	20	10	158	11	---	64	67	16	54
Lys	50	297	67	48	6	10	16	25	29	17	8	29	12	13	103	32	---	134	30	13
Arg	35	227	23	12	35	18	46	46	18	27	23	13	18	22	41	13	99	---	79	37
His	40	161	14	13	0	19	18	3	0	1	1	8	2	0	20	1	71	28	---	43
From	38	112	10	47	0	7	0	10	69	2	24	6	21	47	51	21	17	6	12	---

You can see the frequency table I showed this earlier so, you can see the some mutations are highly preferred because, they experimental lists they do not want to spend the

money for a any on order mutant. So, they like to have some specific mutations to understand, what will happen with their particular mutation, this is why if you see many mutations to alanine to understand the effect of this is a side chain right.

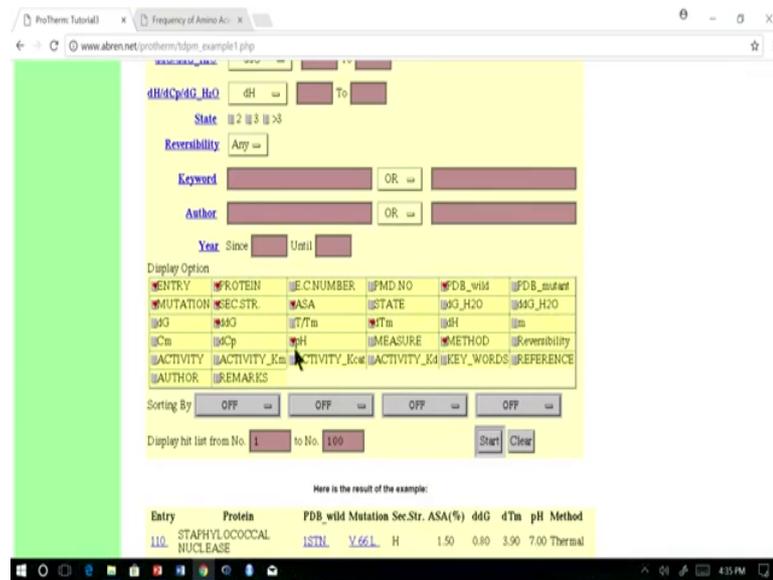
So, likewise you can get the statistics right. Now, if you go to the tutorial right. So, if it is a so, here we gave four different a questions first question is search a data for single mutation, in helical sequence obtained with thermal denaturation experiments, what will do with this what are the search after you need to use right.

(Refer Slide Time: 21:49)



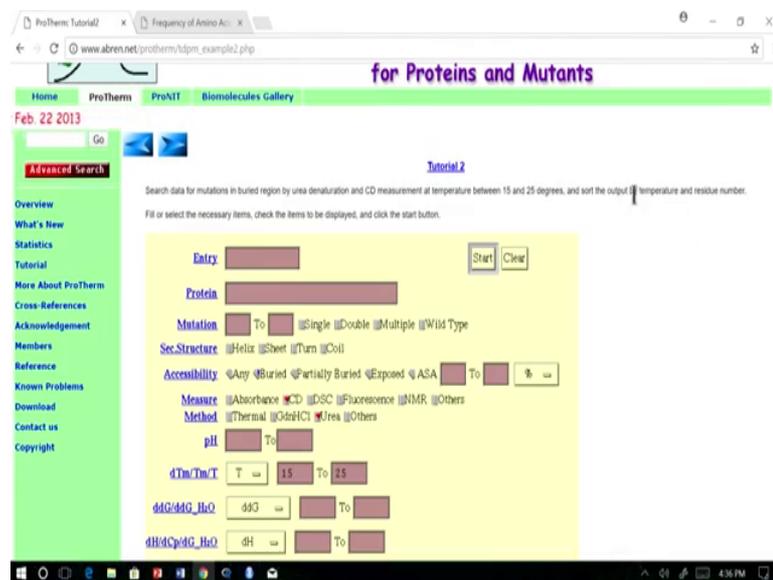
We need the because the question is helical sequence. So, we need to a click on a helical and the single mutation and, you can click on the single mutation right. And the accessibility does not matter because; we wrote we are not interested in we taking anything on the pH 7. So, you have got the pH 7 to 7 thermal denaturation. So, you put thermal now if you do like this.

(Refer Slide Time: 22:13)



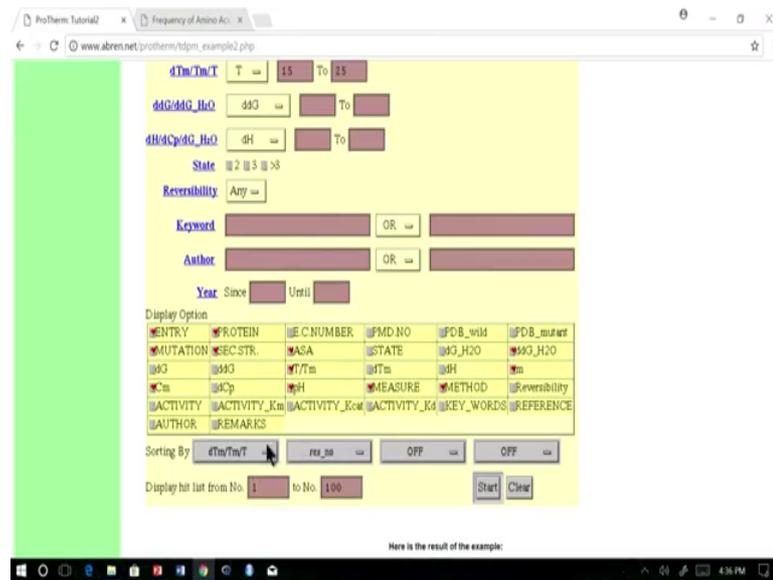
Right then, currently we click the display option, then we will start this then finally, we get the data now it is symbol the first one.

(Refer Slide Time: 22:24)



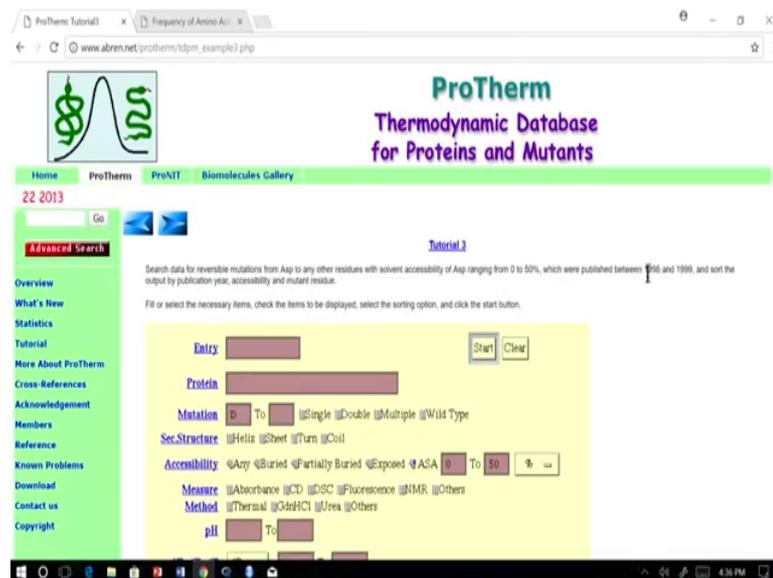
And go a second one this is with the denaturant denaturation here, we have the buried helix region, buried means, you have to get the accessibility buried right and, then you can see the urea denaturation. So, you put the urea and the CD measurement you put the CD and the temperature between 15 and 25 degrees, you see the temperature and, if we start out by temperature residue number.

(Refer Slide Time: 22:46)



So, in this case sorting you need to change temperature residue number right and, if we start and we will get the data is the one for these things right. Then the will take one more example.

(Refer Slide Time: 23:01)



So, for example, reversible mutation and aspartic acid to any other residues; so, wild type residues aspartic acids now here is aspartic residues here, D to any other residue. So, in this case I put blank right, then you use the other conditions accessibility and the year of publication between 96 to 99 right.

(Refer Slide Time: 23:21)

The screenshot shows the ProTherm search interface with the following search criteria:

- Accessability:** Buried, Partially Buried, Exposed, ASA (0 to 50)
- Measure:** Absorbance, CD, DSC, Fluorescence, NMR, Others
- Method:** Thermal, GdHCl, Urea, Others
- pH:** () To ()
- dTm/Tm/T:** dTm () To ()
- dG/dG_H2O:** dG () To ()
- dH/dG_H2O:** dH () To ()
- State:** 2, 3, >3
- Reversibility:** Yes
- Keyword:** () OR ()
- Author:** () OR ()
- Year:** Since 1996 Until 1999

Display Option table:

<input checked="" type="checkbox"/>	ENTRY	<input checked="" type="checkbox"/>	PROTEIN	<input checked="" type="checkbox"/>	EC NUMBER	<input checked="" type="checkbox"/>	PMD NO	<input checked="" type="checkbox"/>	PDB_wild	<input checked="" type="checkbox"/>	PDB_mutant
<input checked="" type="checkbox"/>	MUTATION	<input checked="" type="checkbox"/>	SEC.STR.	<input checked="" type="checkbox"/>	ASA	<input checked="" type="checkbox"/>	STATE	<input checked="" type="checkbox"/>	dG_H2O	<input checked="" type="checkbox"/>	dG_H2O
<input checked="" type="checkbox"/>	dG	<input checked="" type="checkbox"/>	dG	<input checked="" type="checkbox"/>	dTm	<input checked="" type="checkbox"/>	dTm	<input checked="" type="checkbox"/>	dH	<input checked="" type="checkbox"/>	dH
<input checked="" type="checkbox"/>	MEASURE	<input checked="" type="checkbox"/>	METHOD	<input checked="" type="checkbox"/>	REVERSIBILITY	<input checked="" type="checkbox"/>	KEYWORDS	<input checked="" type="checkbox"/>	REFERENCE	<input checked="" type="checkbox"/>	REMARKS

So, if you used all these things 96 and 96 and you will get the complete data right.

(Refer Slide Time: 23:25)

The screenshot shows the search results page with the following table:

Entry	PDB_wild	Mutation	Sec.Str.	ASA(%)	dG_H2O	Tm	pH	Method	Reversibility	REFERENCE
5716	IRNL	D.76.A	S	2.36	NULL	15.00	7.00	Urea	YES	BIOCHEMISTRY.38.13379-13384(1992).
5720	IRNL	D.76.A	S	2.36	NULL	35.60	7.00	Thermal	YES	BIOCHEMISTRY.38.13379-13384(1992).
5714	IRNL	D.76.N	S	2.36	NULL	15.00	7.00	Urea	YES	BIOCHEMISTRY.38.13379-13384(1992).
5718	IRNL	D.76.N	S	2.36	NULL	37.00	7.00	Thermal	YES	BIOCHEMISTRY.38.13379-13384(1992).
5726	IRNL	D.76.N	S	2.36	2.58	25.00	7.00	Urea	YES	BIOCHEMISTRY.38.13379-13384(1992).
5727	IRNL	D.76.N	S	2.36	3.53	20.00	7.00	Urea	YES	BIOCHEMISTRY.38.13379-13384(1992).
5728	IRNL	D.76.N	S	2.36	4.25	15.00	7.00	Urea	YES	BIOCHEMISTRY.38.13379-13384(1992).
5729	IRNL	D.76.N	S	2.36	5.00	10.00	7.00	Urea	YES	BIOCHEMISTRY.38.13379-13384(1992).
5730	IRNL	D.76.N	S	2.36	5.55	5.00	7.00	Urea	YES	BIOCHEMISTRY.38.13379-13384(1992).

So, you have various options in ProTherm and you can use the multiple combination of different options to get your information fine. So, we use various terms right so, what is the meaning of accessibility and for example, I will put the buried and partially buried and exposed, what is the meaning of exposed.

Student: (Refer Time: 23:45).

Accessibility service is very high it is outside the protein. So, what did they range we use for exposed, we do not know right.

Student: (Refer Time: 23:53).

(Refer Slide Time: 23:53)

Accessible Surface Area	Accessible surface area (ASA) of the residue in wild type (computed using the program ASC, Analytical Surface Calculation) expressed in units of Å ² . Accessibility (%) is defined as the ASA of the residue at the mutation site (X) in its parent protein, computed with ASC divided by the ASA of the residue in an extended tripeptide Ala-X-Ala conformation. The extended state ASA was calculated using ECEPP-2 algorithm with dihedral angles given by Oobatake and Ooi (Prog. Biophys. Mol. Biol. 1993) 59, 237-284) and the van der Waals radius of atoms from Ooi et al. (Proc. Natl. Acad. Sci. USA (1987) 84, 3086-3090). The values are Ala-110.2; Asp-144.1; Cys-140.4; Glu-174.7; Phe-200.7; Gly-78.7; His-181.9; Ile-185.0; Lys-205.7; Leu-183.1; Met-200.1; Asn-146.4; Pro-141.9; Gln-178.6; Arg-229.0; Ser-117.2; Thr-138.7; Val-153.7; Trp-240.5; Tyr-213.7 (the units are in Å ²). We classified the residues with less than 20% accessibility as buried, between 20% and 50% as partially buried and more than 50% as exposed.
Measure	The experiments performed to measure the thermodynamic parameters (Fluorescence spectroscopy, Circular Dichroism, Differential Scanning Calorimetry, Absorbance, NMR, etc.) Keywords: FL, CD, DSC, Abs, NMR, others
Method	Experimental method of denaturation (keywords: Thermal, Urea, GdnHCl etc.). (activity: 50% relative remaining activity of the enzyme after the heat treatment)
pH	the pH value.
m	Slope of dG on denaturant concentration (dG vs urea GdnHCl, dG = dG(H ₂ O) - m[Denaturant]). Unit is kcal/mol M
Cm	Concentration of denaturant at which 50% of the protein is unfolded [M]
dTm	[Tm(mutant) - Tm(wild)] [degree Celsius]
T	In the case of denaturant denaturation methods, T is the temperature used in the experiment. In the case of thermal denaturation methods, T is the temperature at which thermodynamic quantities such as unfolding free energy (ΔG), unfolding free-energy change due to mutation (ΔΔG) and sometimes unfolding enthalpy change (ΔH) were calculated [degree Celsius]
Tm	Midpoint temperature of the thermal unfolding for thermal denaturation methods [degree Celsius]
ΔG	1) Free energy of unfolding at a certain concentration of denaturant in the case of denaturant denaturation methods 2) Free energy of unfolding obtained for extrapolation temperature T using dCp in the case of thermal denaturation method [kcal/mol]
ΔΔG	ΔG(mutant) - ΔG(wild) [kcal/mol] Free energy of unfolding obtained with Schellman equation (ΔΔG = dTm dS) in the case of thermal denaturation method [kcal/mol]
ΔΔG_H2O	ΔG_H2O(mutant) - ΔG_H2O(wild) [kcal/mol]
ΔH	Enthalpy change of denaturation [kcal/mol]
ΔH_H	van't Hoff enthalpy change of denaturation (enthalpy obtained from the temperature dependence of the denaturation equilibrium constant)
ΔHcal	Calorimetric enthalpy change of denaturation (enthalpy measured by calorimetry)

This will tell you what are the conditions, we used for this ProTherm, because we can use different cutoff values for the buried or partially exposed. In this case we use less than 20 percent buried that is very wide, and 20 to 50 as partially buried and more than will be exposed.

If you want to restrict this 20 percent as 5 percent what will you do, can we get it get the data because here the buried we define as to 0 to 20 percent. If you are interested to get 0 to 5 can we get the get it that you get the data.

Student: Yeah.

Yes right. So, here you can define instead of buried, you have to give the value 0 to 5, now 0 to whatever you like. So, then you are also get the data right.

(Refer Slide Time: 24:39)

Terms	Explanations
No	Entry number. This option can be used for getting data from a particular entry (Eg. 3012) or a range of entries (Eg. 10107-10365) with/without other search conditions
Protein Name	Name of the protein. Multiple words can be entered with spaces. Wild card can also be used. * for a string of characters and ? for a character. If * and ? were to be used as real characters, place "" (backslash) before them
E.C.No.	Enzyme Commission number
PMD No.	Protein Mutant Database accession number
PDB_wild	Protein Data Bank code for the native protein
PDB_mutant	Protein Data Bank code for the mutant protein
STATE	Number of transition states
Mutation	Details about the mutation: residue in wild type, residue number and residue in mutant protein (e.g. A 123 G). In the case of insertion or deletion mutations, all inserted/deleted residues appear along with the preceding residue that corresponds to the residue number at which the mutation occurs (for example, A 378 AVL refers that the two amino acid residues "V" and "L" are inserted at 378th position where the residue "A" is present in the wild type, and it is represented conversely for the deletion, viz. AVL 378 A). The mutation residue number is given as in the reference. In some cases, the residue number may not match with PIR and PDB sequences. In such cases, the residue numbers corresponding to PIR and PDB sequences are also given in the parenthesis.
Secondary Structure	Secondary Structural Information for the mutation site (Helix, Strand, Turn and Coil; we obtain the data from PDB) In search result: H = Helix, S = Strand, T = Turn, C = Coil
Accessible Surface Area	Accessible surface area (ASA) of the residue in wild type (computed using the program ASC, Analytical Surface Calculation) expressed in units of Å ² . Accessibility (P _{sa}) is defined as the ASA of the residue at the mutation site (X) in its parent protein, computed with ASC divided by the ASA of the residue in an extended tripeptide Ala-X-Ala conformation. The extended state ASA was calculated using ICEPP-2 algorithm with dihedral angles given by Dobson and Ooi (Prog. Biophys. Mol. Biol. 1993) 59: 237-284) and the van der Waals radius of atoms from Ooi et al. (Proc. Natl. Acad. Sci. USA, 1987) 84: 3086-3090). The values are Ala-110.2; Asp-144.1; Cys-140.4; Glu-174.7; Phe-200.7; Gly-78.7; His-181.9; Ile-185.0; Lys-205.7; Leu-183.1; Met-200.1; Asn-146.4; Pro-141.9; Gln-178.6; Arg-229.0; Ser-117.2; Thr-138.7; Val-153.7; Trp-240.5; Tyr-213.7 (the units are in Å ²). We classified the residues with less than 20% accessibility as buried, between 20% and 50% as partially buried and more than 50% as exposed.

So, here the help H, you can get the information for all the technical terms right, we are used in the ProTherm database. So, you see all the technical terms and, they all the terms you have the a complete information.

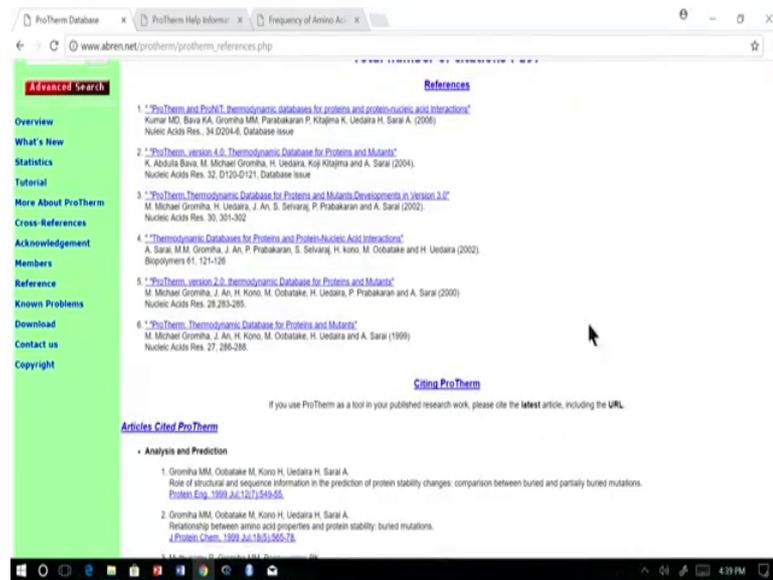
(Refer Slide Time: 24:48)

Activity_Kd	Dissociation constant [nM]
Author	Name of the authors. Multiple words can be entered with spaces. Wild card can also be used. * for a string of characters and ? for a character. If * and ? were to be used as real characters, place "" (backslash) before them. These can be combined with AND and OR logical operations to form more complex patterns.
Year	Year of publication.
Reference	Complete reference of the article with a link to NCBI database with PMID
Remarks	Some specific comments
Sorting By	Sorting option. Check the items to sort the list.
Source	Source of the protein
Molecular weight	Molecular weights were calculated as follows: The amino acid sequence corresponding to a protein was taken from PIR database, and leader (signal) regions were removed. Then, the molecular weights were calculated according to the amino acid sequence. Amino and carboxy terminals were taken into account. Any mutations to the amino acid sequence were taken into account. However, post-translational modifications were NOT taken into account.
Length	Total number of Amino Acid residues in the protein
No of molecule	Number of molecule (e.g. 1 = Monomer, 2 = Dimer, etc.)
Buffer_name	Name of the buffer used in the experiment
Buffer_conc	Concentration of the buffer
Ion_name	Name of the added ion
Ion_conc	Concentration of the ion
ADDITIVES	Details about the additives (e.g. glycerol)
Protein_conc	Concentration of the protein when the experiment has been performed
PIR_ID	Protein Information Resources Codes
SWISS_PROT ID	Protein Sequence Knowledge Base Codes
RELATED_ENTRIES	List of all the entries that contain other data reported in the same reference.

Now, in this case you can understand more about the terms used in the ProTherm database right.

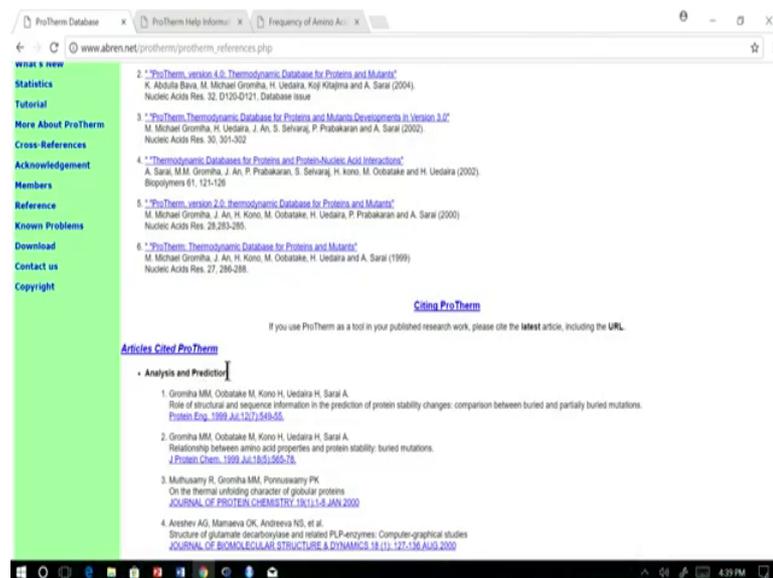
So, what are the applications of ProTherm. So, for what we use ProTherm there is no recent updates but if you go to the references ok, reference if you see.

(Refer Slide Time: 25:15)



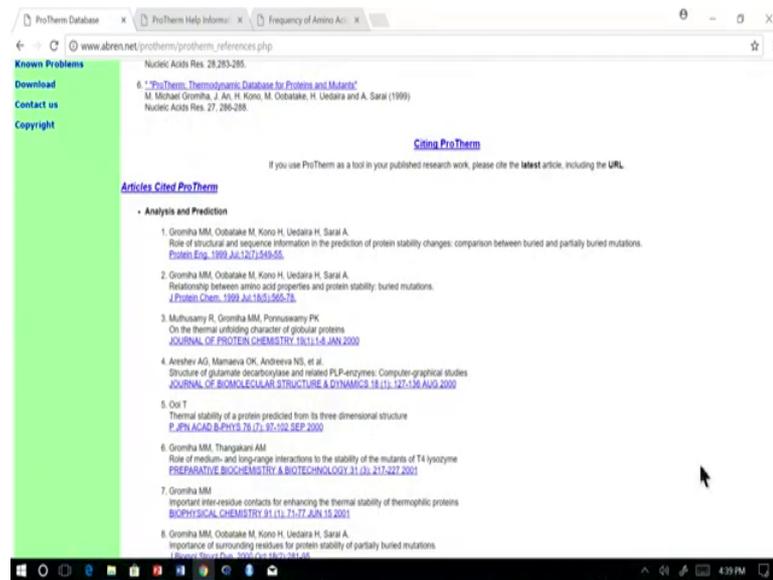
So, the various application these are the database papers.

(Refer Slide Time: 25:19)



And we can do the analysis and prediction, you can collect the data and, we can analyze the stability data, obtained from the denaturant or thermal denaturation, with the respect to accessibility surface area or the pH or the external conditions, or the secondary structures. And we can see how the stability varies, what happened the stability in different secondary structure and you can do the analysis.

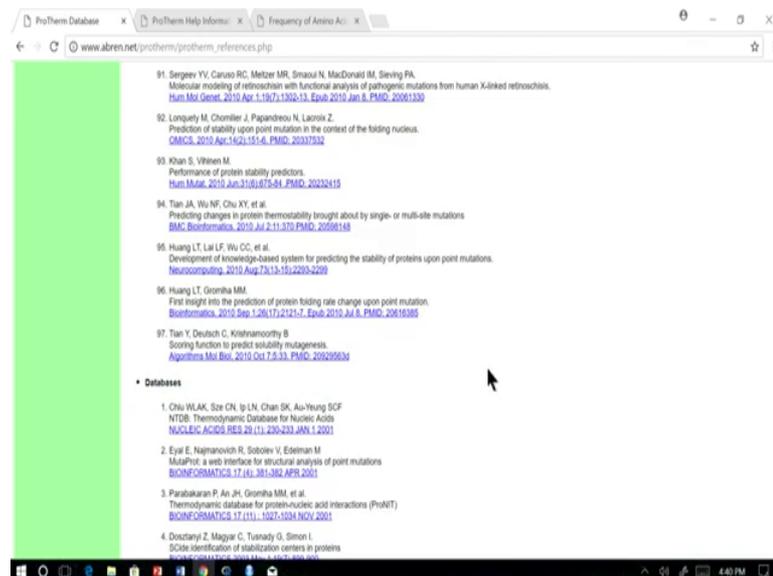
(Refer Slide Time: 25:41)



What we do the analysis then you can also predict, we can develop several models regression the models, or the machine learning techniques right, or the knowledge based predictions, or the decision tree models you can use various models to predict the stability upon mutation.

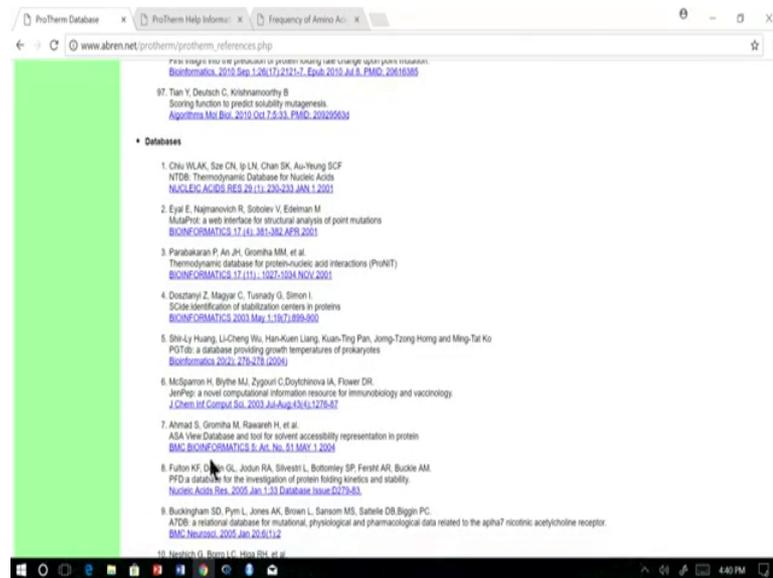
Then you can also get some a verify with the experimental data for example, if there are some prediction methods they predict some data, these are the various citation currently we are not updated is ok.

(Refer Slide Time: 26:06)



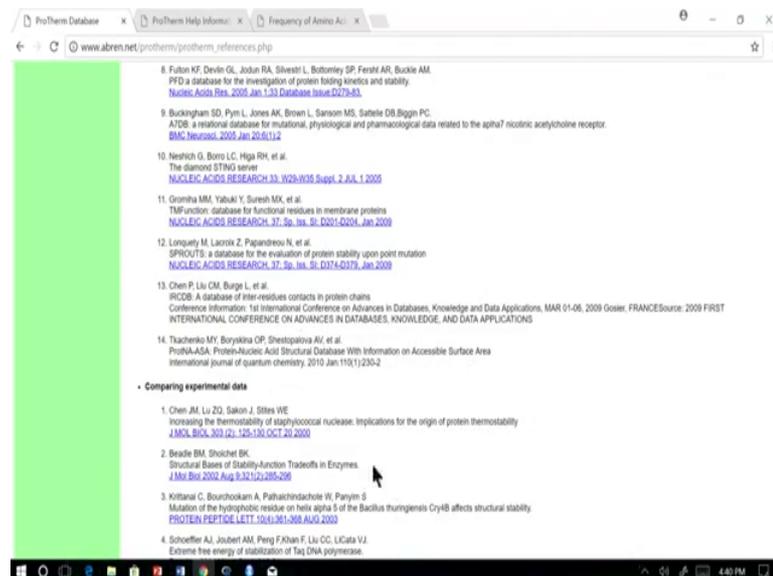
So, then it is used to cross link with our databases because, ProTherm is linked with the PDB and also various other databases, we have more than 25000 data. So, these are the various DDB and the pronate slide P GDB.

(Refer Slide Time: 26:24)



All these databases they linked these ProTherm database. So, we have the structural information and, we can related it with the thermodynamics. So, we can see to understand the structure stability information with the function and diseases. In this case this database is a very useful. So, there are various databases.

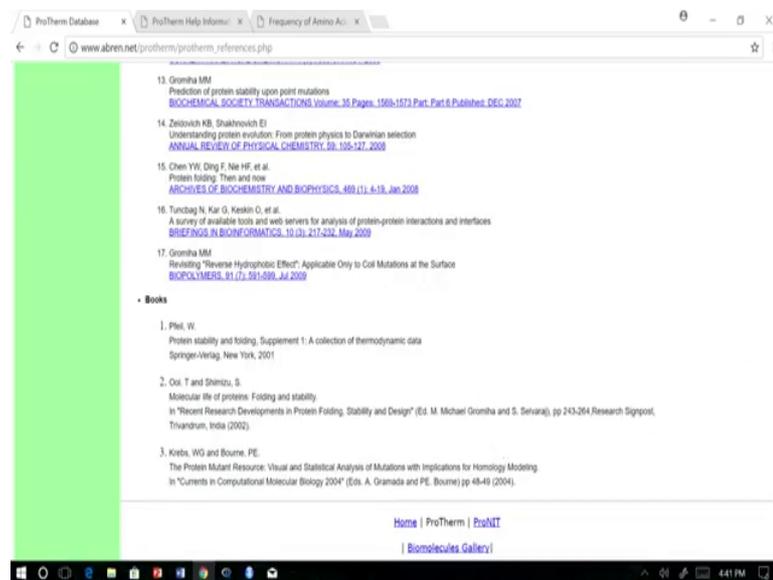
(Refer Slide Time: 26:45)



Then any methods so, worked out they can predict, or they can finally, explain some reasons for any mutations, we can verify with the datas available in the ProTherm or not, this experiment only available experimental database. So, we can compare whether this is comparable with the other therm mutations in the ProTherm database or not right this is that, they use.

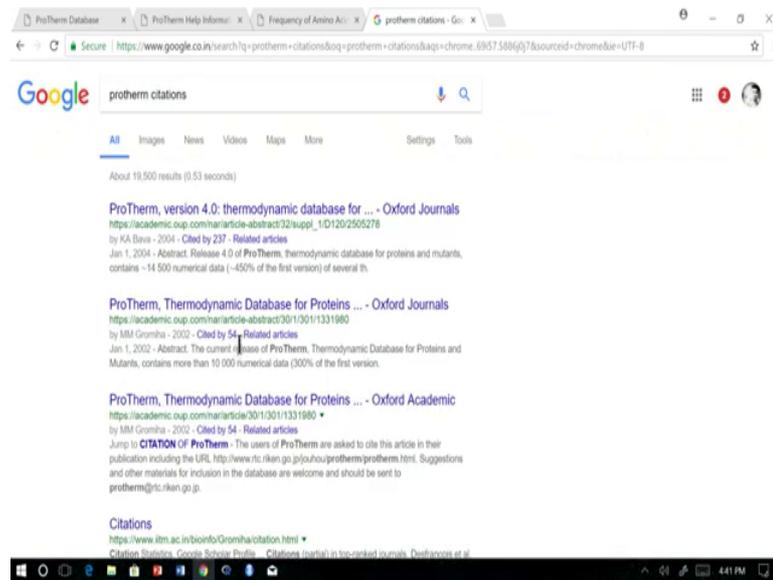
Now, there are many many citations, then also cited and very general applications and why are we can receive for the by technological applications, with increase in stability. So, right they are used the ProTherm database. So, there is several review article also they all sent the books.

(Refer Slide Time: 27:21)



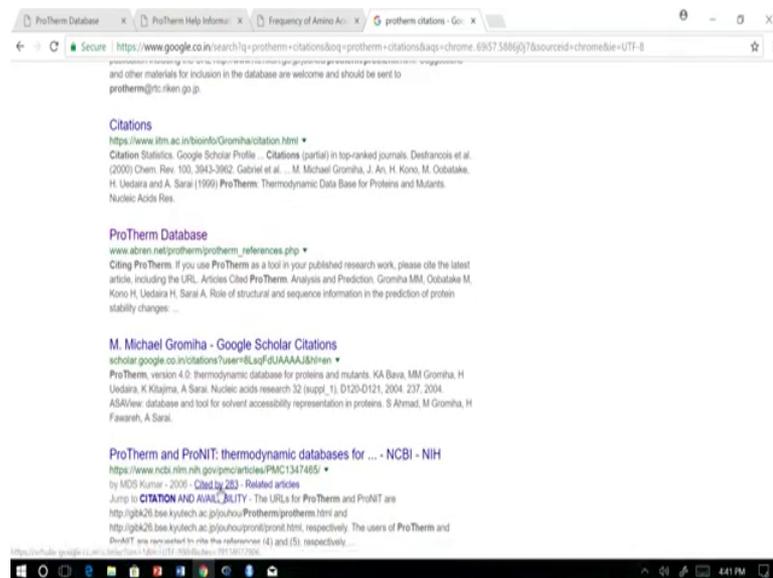
So, they also cite the ProTherm. So, this will tell the how the ProTherm is a important in this case.

(Refer Slide Time: 27:30)



So, if you search the ProTherm citations, you will get yeah this is the ProTherm version 4 resided with 237 times, these 54 this is 54 right there are many citations right ok.

(Refer Slide Time: 27:50)



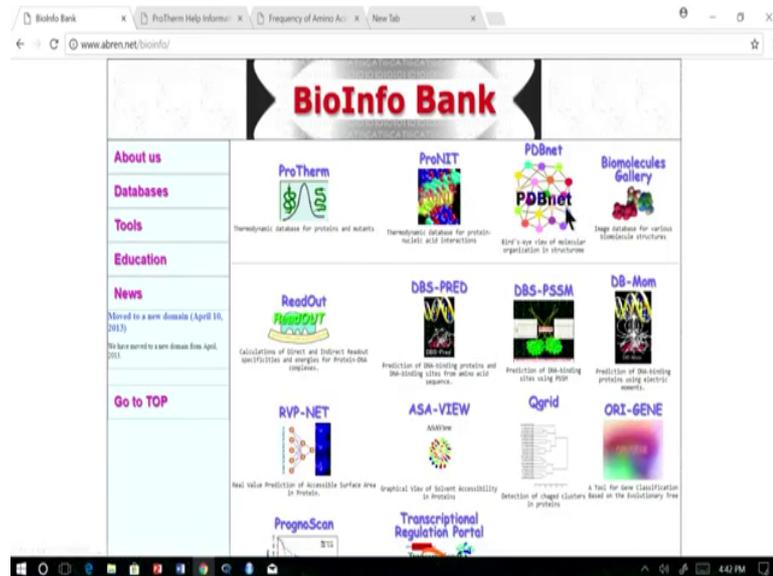
This is 283 citations to some citations, it totally in a if you add up all these 4 5 papers right.

So, we get more than 1000 citations, because earlier years nucleic acid research they published the updates of this in databases right not currently earlier days. So, we have various updates, we have 4 or 5 updates. So, there is a recent we have the different

updates, but each version we introduce new features as well as a increase the number of data and so, this is the reason why it case.

So, it is only a used resource for protein stability.

(Refer Slide Time: 28:28)



So, it is a good database and useful resource right for the protein researches, these are other databases available in this bank. So, related with the ProTherm as well as other things. So, you can use the ProTherm database and, get the information and have fun with this data and try to explore right more details and more insides about protein stability and relate with the functions, or the diseases and that right.

Thank you.