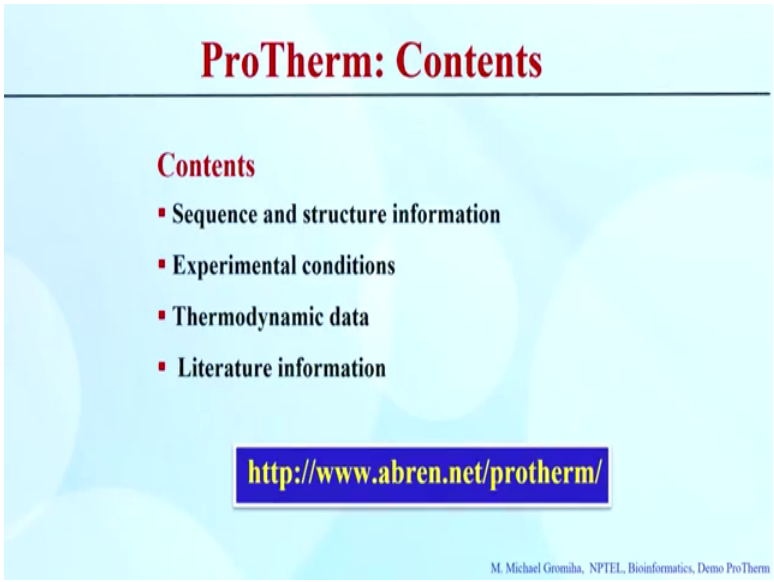
Bioinformatics
Dr. M. Michael Gromiha
Department of Biotechnology
Indian Institute of Technology, Madras

Lecture – 49
Demonstration on ProTherm: Thermodynamic database for proteins and mutants

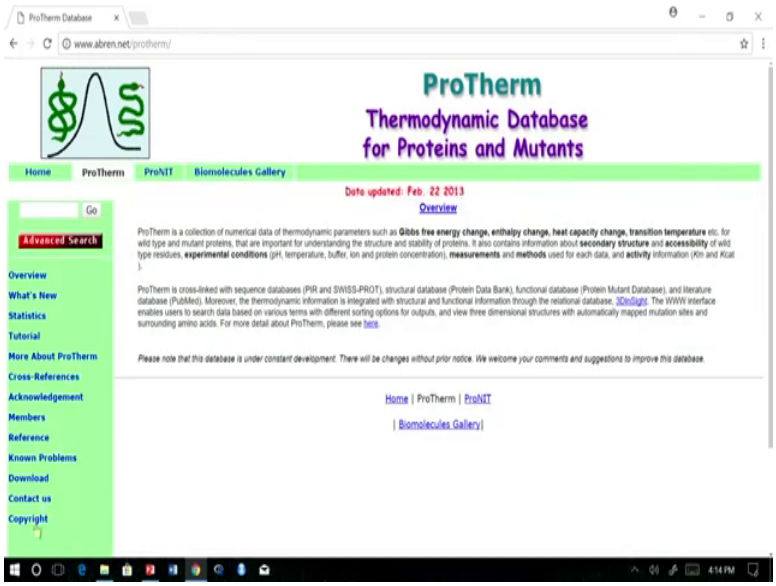
Demonstration on ProTherm is Thermodynamic Database for proteins and mutants.

(Refer Slide Time: 00:28)



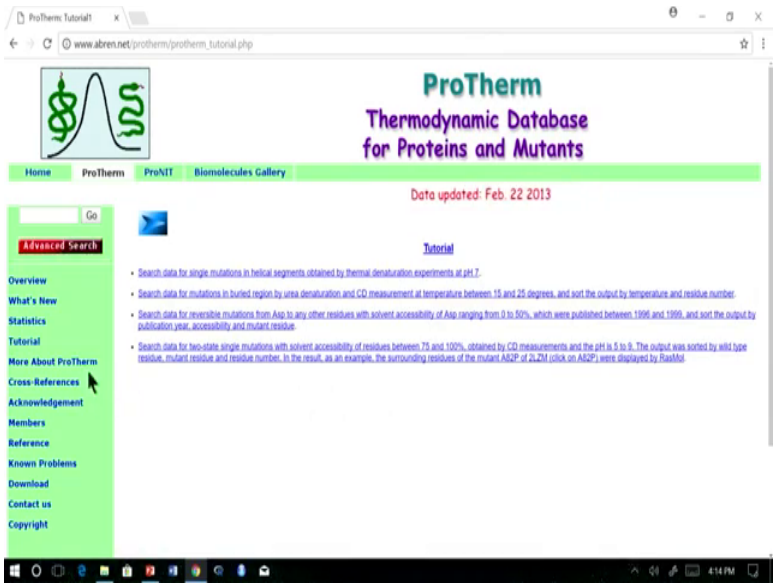
As we discussed earlier ProTherm contents information on sequence instruction information, experimental conditions thermodynamic data and so, literature information for different proteins, as well as their mutants. The major aspect of ProTherm is thermodynamic data and supplemented with all the other 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)



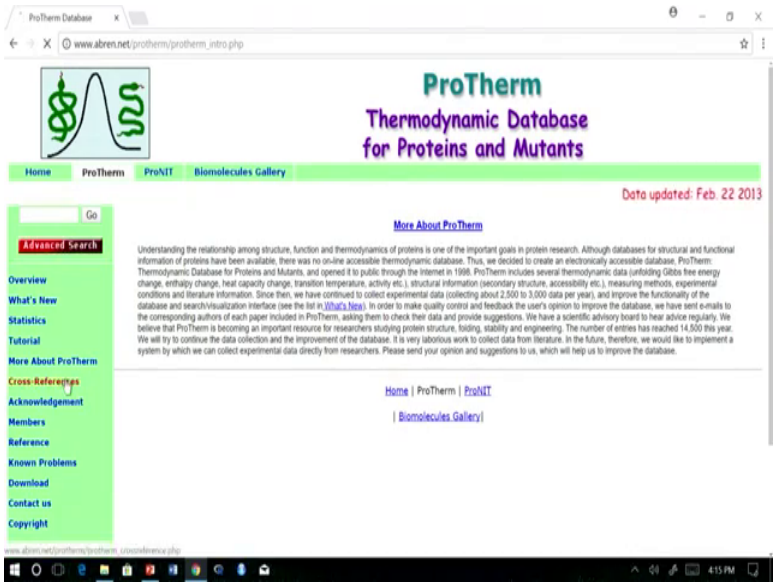
Oh this is website for ProTherm. So, here you can see the links with the other databases like pronate or biomolecules gallery, and left side, you can see the major aspects of ProTherm, you can see the overview. So, what the major contents of this a database and, what we can obtain from this database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

(Refer Slide Time: 01:25)



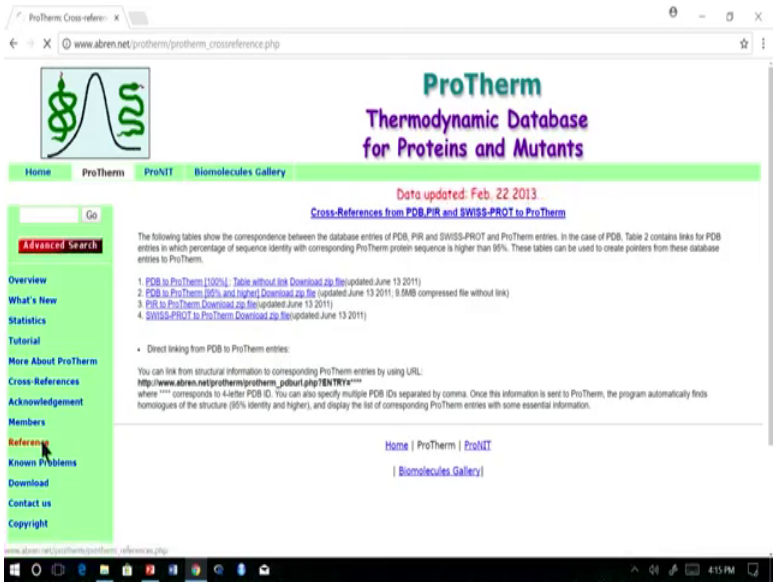
So, here you can see details how to obtain data, or how to retrieve data from ProTherm database and, here you can give the more details about the ProTherm all right.

(Refer Slide Time: 01:36)



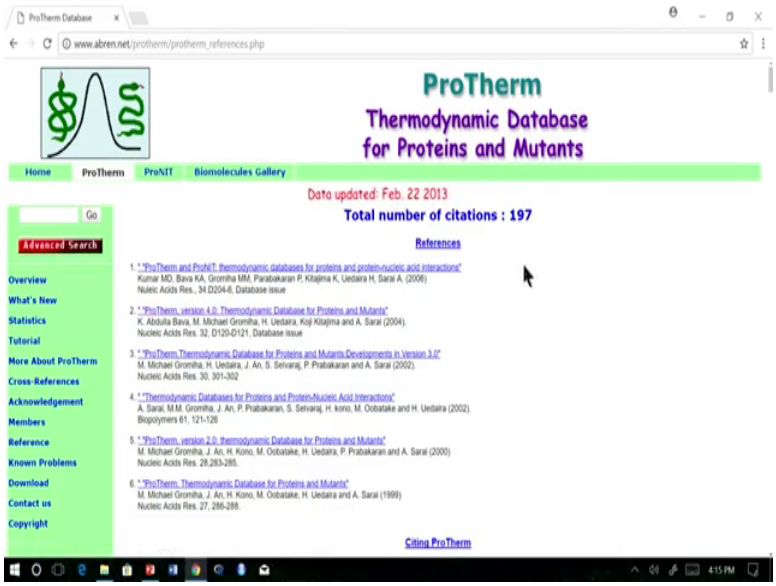
So, now, here is a cross references.

(Refer Slide Time: 01:39)



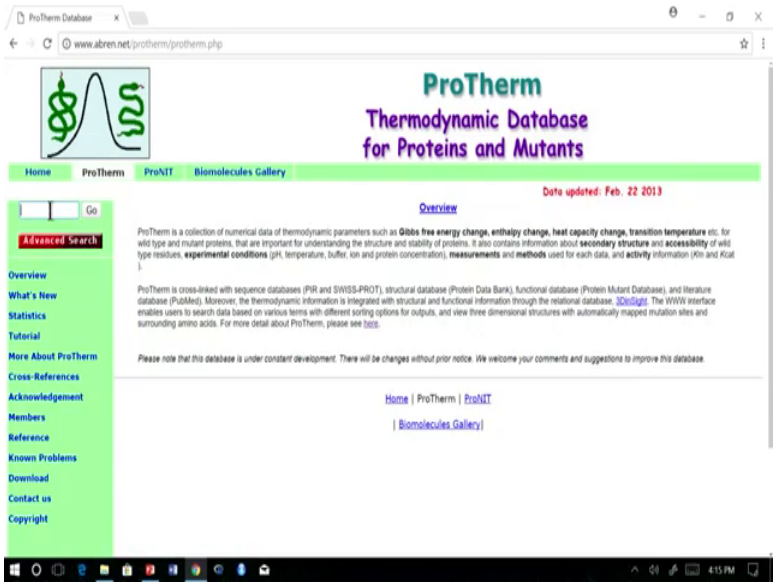
You can see the PDB map to the ProTherm and Swiss proto ProTherm and so, on.

(Refer Slide Time: 01:48)



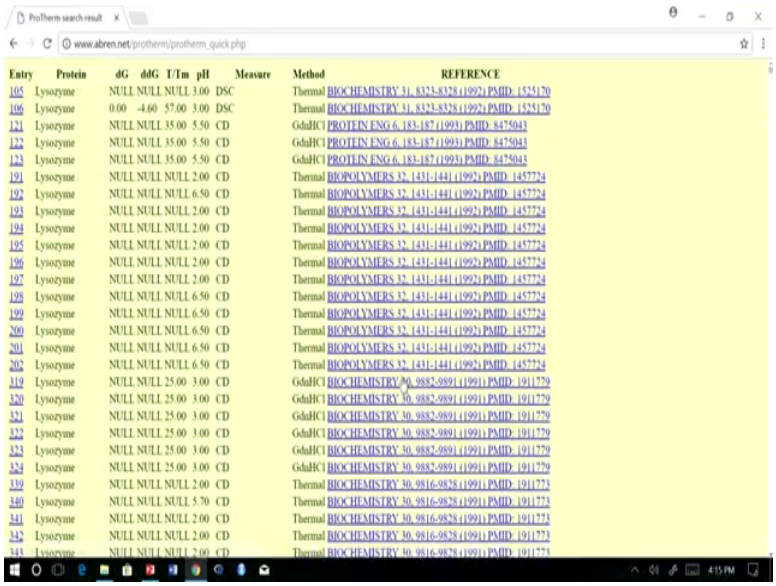
And here we give the references this is the major references and the citations, currently if you look at the different versions of these a papers and website, because more than a 1000 right.

(Refer Slide Time: 02:06)



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.

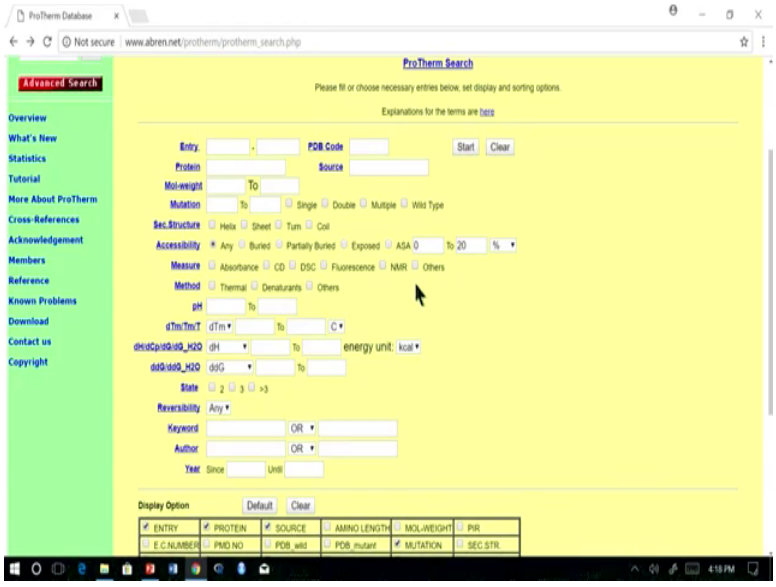
(Refer Slide Time: 02:20)



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)



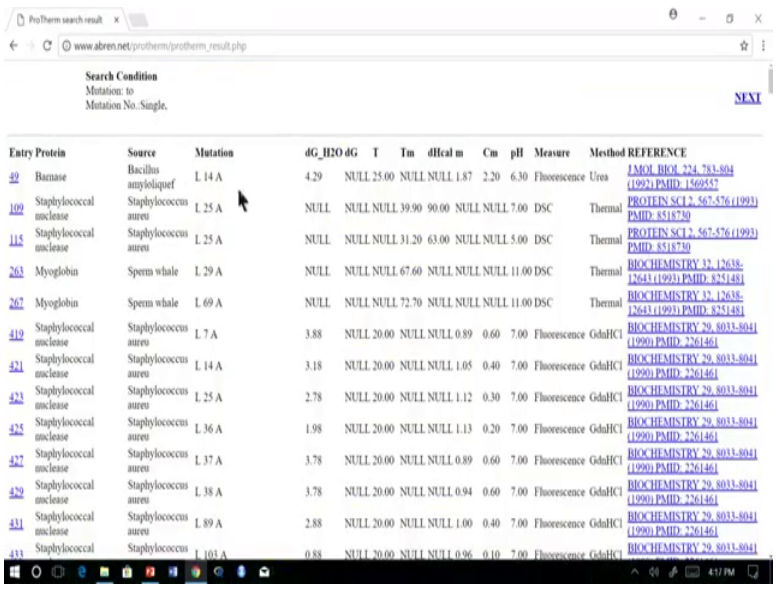
Right here you have a lot of options available to search, you can search with their entry this is the ProTherm entry. So, this is only specific to a ProTherm that, if you have any

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)



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

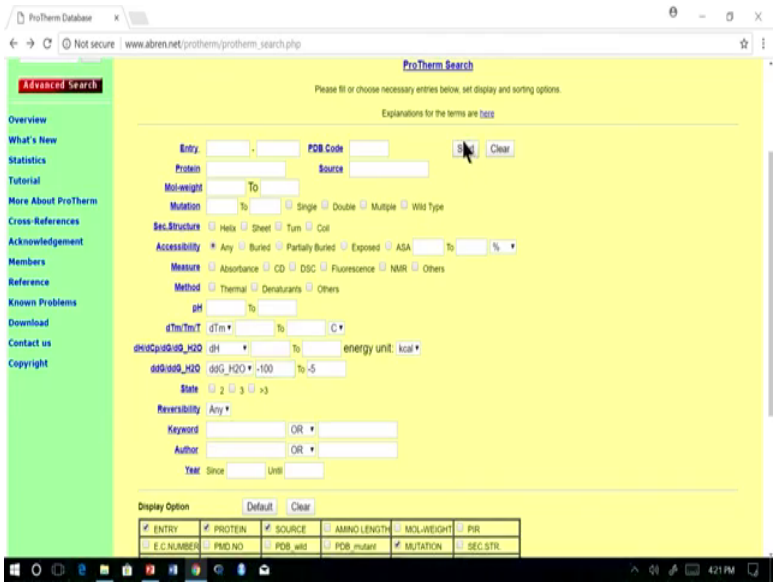
using DSSP, you get I see an X X square, how do we the percentage values. You would do it with the X units, or accessibility and give the percentage and, you can choose any of these options either the directly from the DSSP result, or you can use the percentage you can normalize with the X and X accessibility.

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.

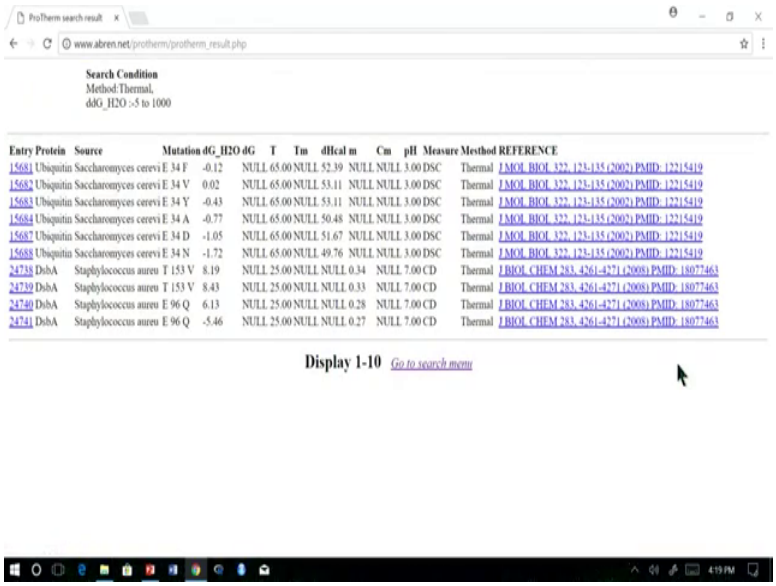
And also you can search for the different ranges of these experimental data for example, if you are interested in the very extreme stability for example, extremely stabilize, or extremely destabilize right. In this case you can give the delta delta G values as more than minus 10 kilo cal per mole and,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



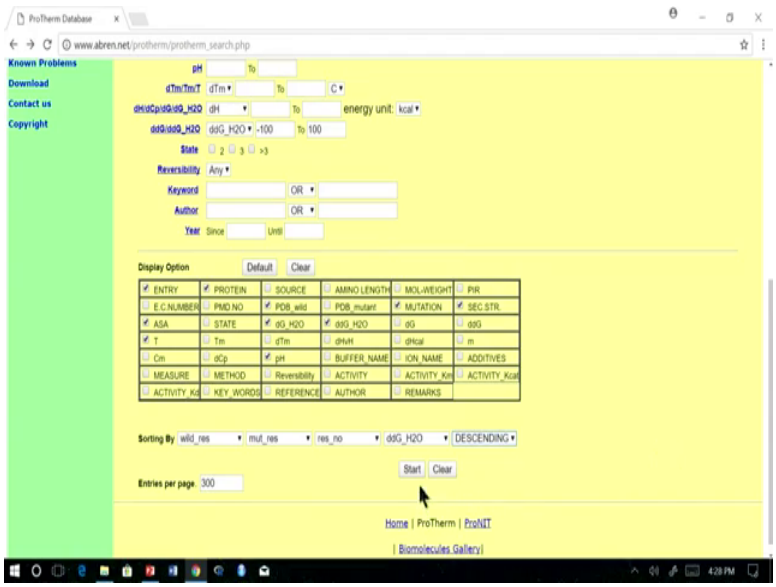
For example you can say you can use it minus 5 all right, then if you see this.

(Refer Slide Time: 05:58)



There are some cases ok; this is minus 5 to any number. So, this what do you get the ok, I will show the delta delta G H 2 O right.

(Refer Slide Time: 06:10)

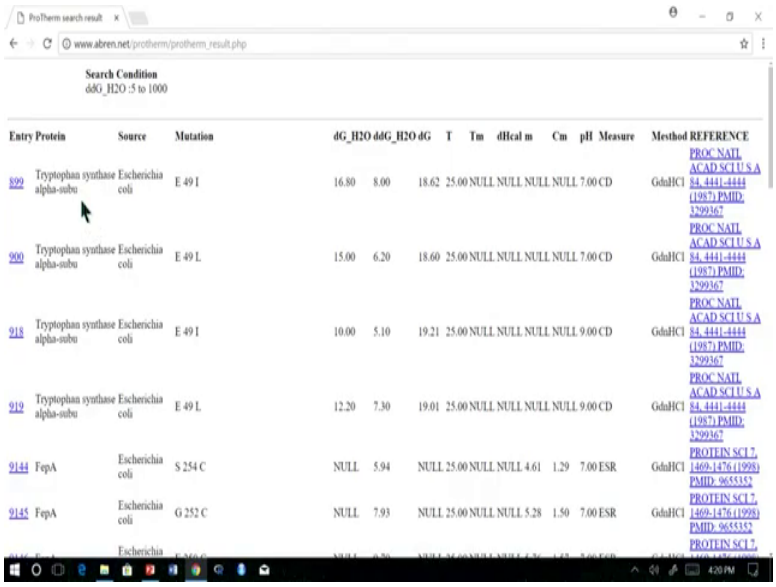


And this is the my the reason is why we choose this what is that because, we could less temperature data, because we put the thermal denaturant. So, right because if you do not get the data with these delta delta G H 2 O only some mutations are matching is a reason.

If you take this out and, if you get the delta delta G is more than 5 because in the ProTherm, we give the positive sign for stabilizing and negative for the destabilizing. So,

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.

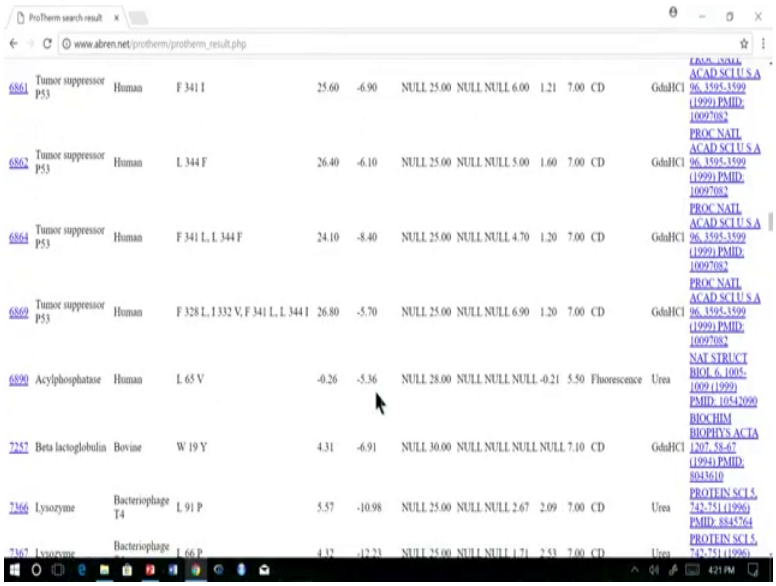
(Refer Slide Time: 07:12)



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)



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.

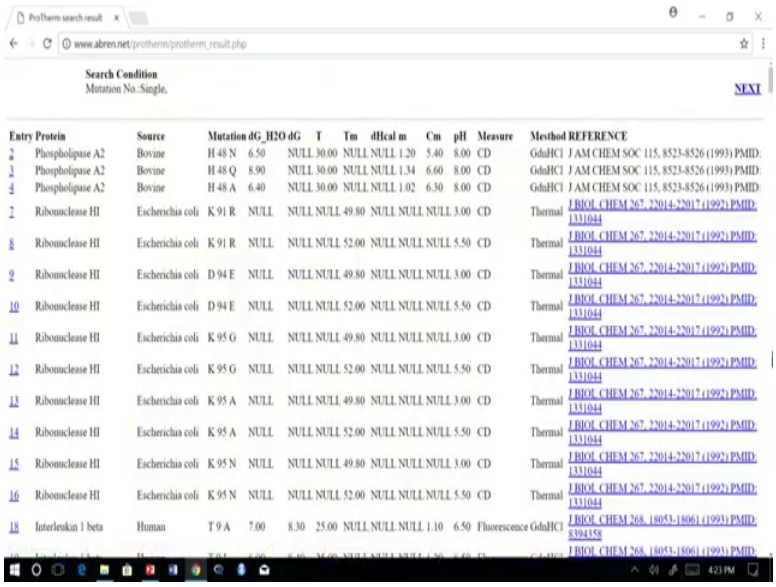
Then you can get reversibility right we can (Refer Time: 08:37) unfolded state and, then you can return back to unfolded state that you say, this is reversibility yes, if not this reversibility is no right like this key words right. So, you can said to the any of the key words as well as the others right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

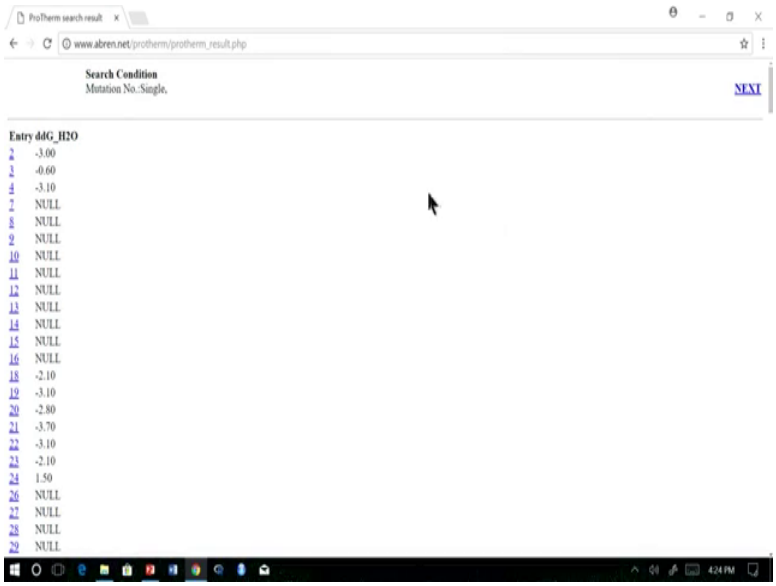


And if you click start now, it is a big mess you need all these things right So, now, the question is among all the information, which share the you are interested in for example, you have a understand disability data because, that is the important aspect of this ProTherm database. So, you want pro thermodynamic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



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).

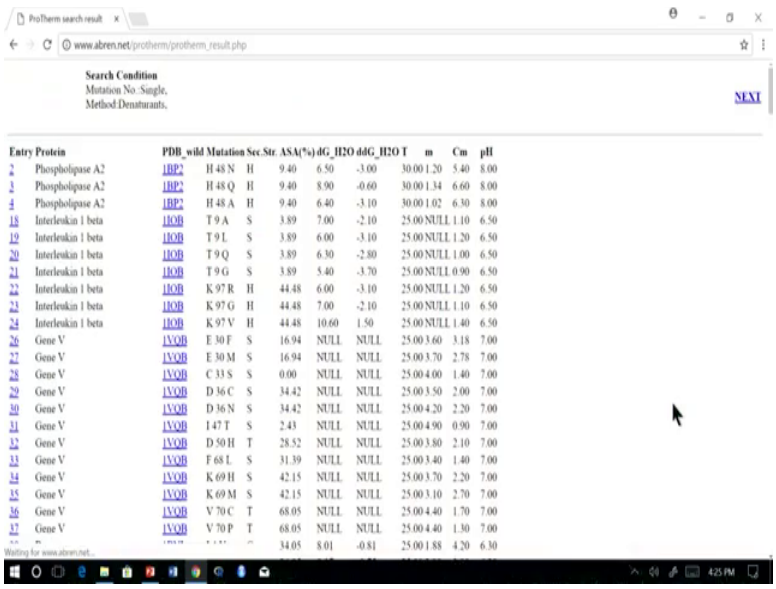
Yeah this is very important right otherwise we do not know the mutation (Refer Time: 11:27) very important that will tell you, work which residue is mutated to which residue in which protein this is fine and anything else do you want.

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_H2O	ddG_H2O	T	m	Cm	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	1IOB	T 9 A	S	3.89	7.00	-2.10	25.00	NULL	1.10	6.50
19	Interleukin 1 beta	1IOB	T 9 L	S	3.89	6.00	-3.10	25.00	NULL	1.20	6.50
20	Interleukin 1 beta	1IOB	T 9 Q	S	3.89	6.30	-2.80	25.00	NULL	1.00	6.50
21	Interleukin 1 beta	1IOB	T 9 G	S	3.89	5.40	-3.70	25.00	NULL	0.90	6.50
22	Interleukin 1 beta	1IOB	K 97 R	H	44.48	6.00	-3.10	25.00	NULL	1.20	6.50
23	Interleukin 1 beta	1IOB	K 97 G	H	44.48	7.00	-2.10	25.00	NULL	1.10	6.50
24	Interleukin 1 beta	1IOB	K 97 V	H	44.48	10.60	1.50	25.00	NULL	1.40	6.50
26	Gene V	1VQB	E 30 F	S	16.94	NULL	NULL	25.00	3.60	3.18	7.00
27	Gene V	1VQB	E 30 M	S	16.94	NULL	NULL	25.00	3.70	2.78	7.00
28	Gene V	1VQB	C 33 S	S	0.00	NULL	NULL	25.00	4.00	1.40	7.00
29	Gene V	1VQB	D 36 C	S	34.42	NULL	NULL	25.00	3.50	2.00	7.00
30	Gene V	1VQB	D 36 N	S	34.42	NULL	NULL	25.00	4.20	2.20	7.00
31	Gene V	1VQB	I 47 T	S	2.43	NULL	NULL	25.00	4.90	0.90	7.00
32	Gene V	1VQB	D 50 H	T	28.52	NULL	NULL	25.00	3.80	2.10	7.00
33	Gene V	1VQB	F 68 L	S	31.39	NULL	NULL	25.00	3.40	1.40	7.00
34	Gene V	1VQB	K 69 H	S	42.15	NULL	NULL	25.00	3.70	2.20	7.00
35	Gene V	1VQB	K 69 M	S	42.15	NULL	NULL	25.00	3.10	2.70	7.00
36	Gene V	1VQB	V 70 C	T	68.05	NULL	NULL	25.00	4.40	1.70	7.00
37	Gene V	1VQB	V 70 P	T	68.05	NULL	NULL	25.00	4.40	1.30	7.00
					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.

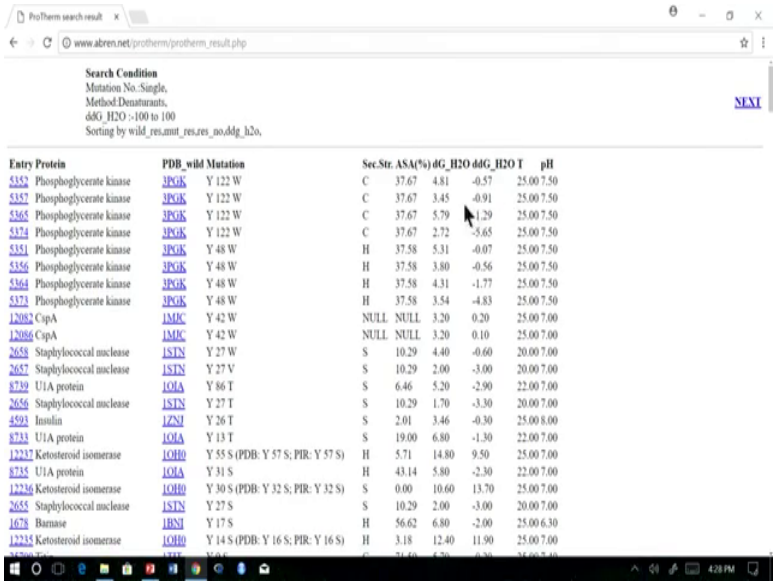
Then in this case also if you see there are many null null data ok, either remove missing AWK comment. Otherwise in the database also you can eliminate these how to eliminate these right, we going to the search option go to delta energy H 2 O, we should get some numbers. So, you can give two extreme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

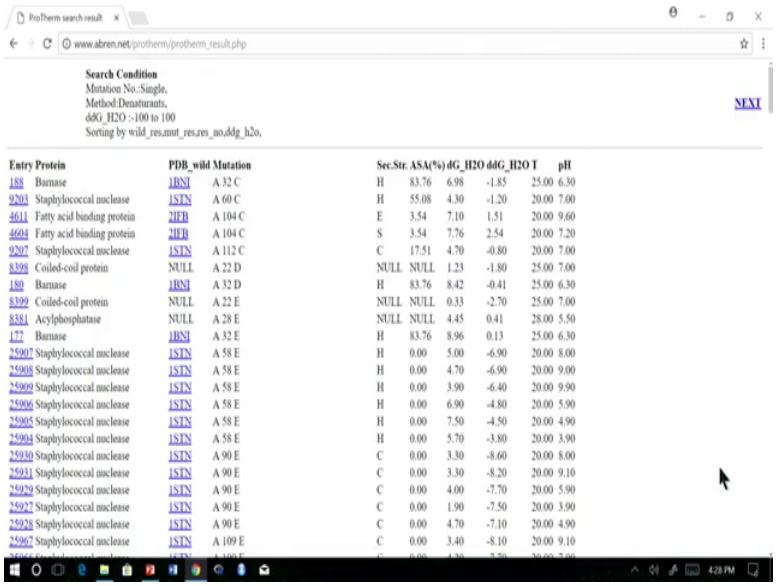
If you do with this wild type residue and the mutant residue and, the residue number ok, then you can also use the values because, delta G H 2 you are now using on so, delta G H 2, then you will see same mutation maybe same protein or different protein and different locations, how we will change right, 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)



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.

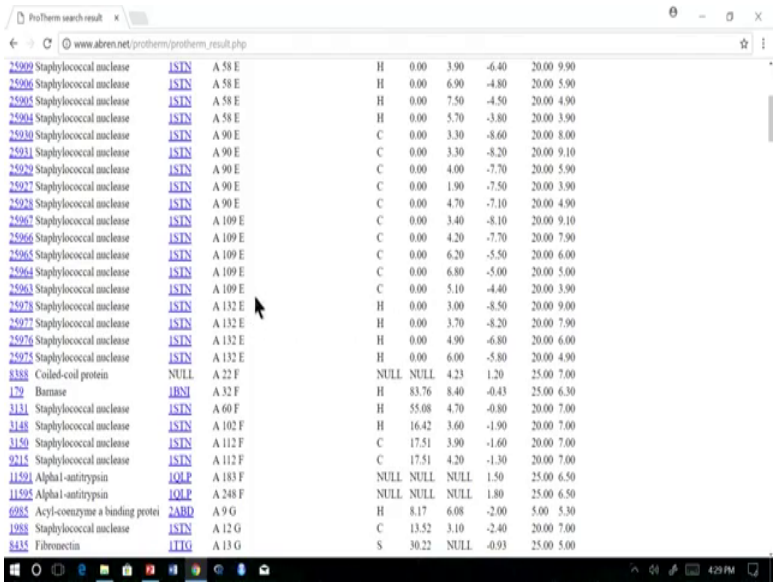
(Refer Slide Time: 15:01)



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.

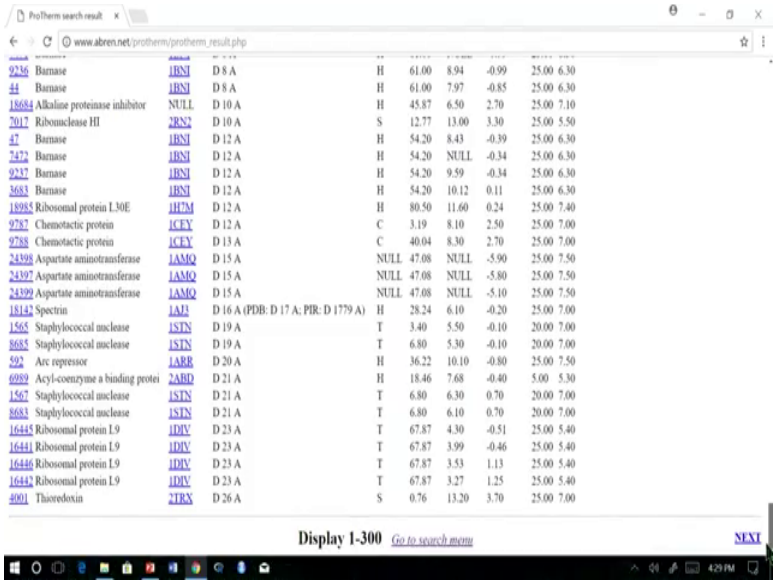
(Refer Slide Time: 15:34)



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.

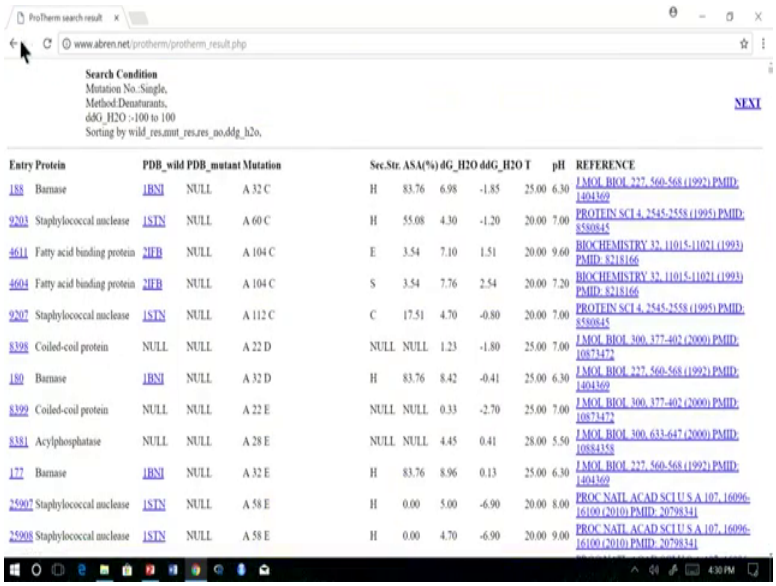
(Refer Slide Time: 16:14)



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)



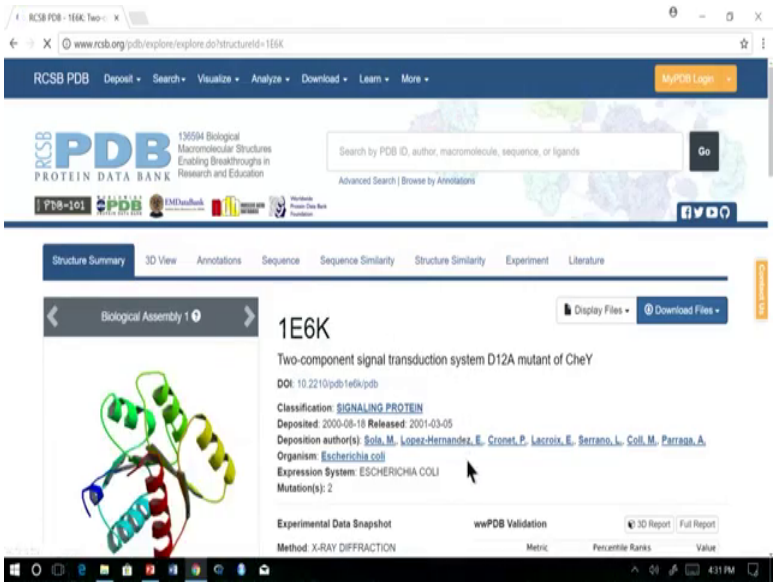
Now, this number is not necessary right. So, you reduce 300 so, we will see these the reference right and the some of the cases PD mutant is most (Refer Time: 17:17) of element if it is lysozyme you can get the a mutant a data.

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

(Refer Slide Time: 17:28)

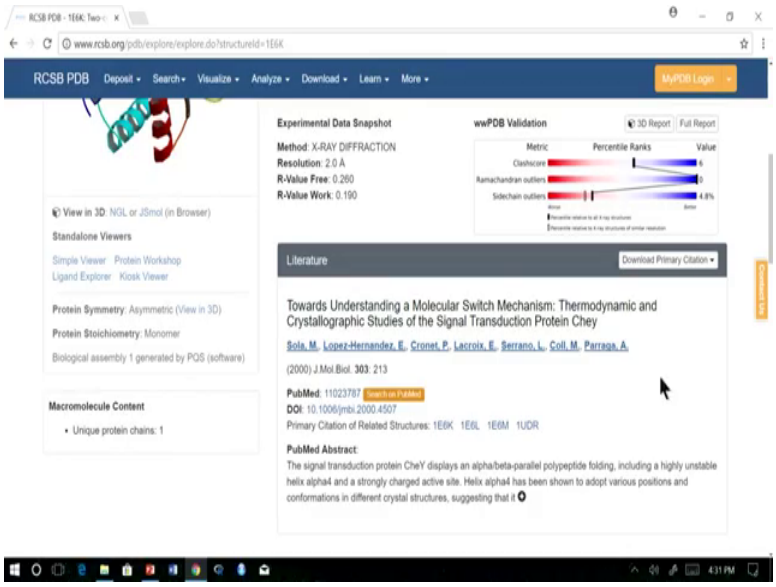
If you click on this one E 6 L here a 1 E 6 k it is a pretty bady.

(Refer Slide Time: 17:36)



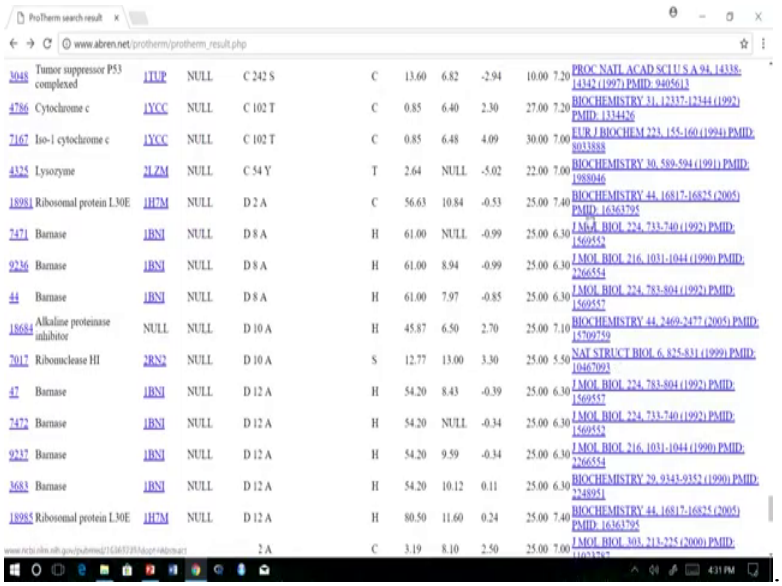
So, you can see this is the mutant all right.

(Refer Slide Time: 17:41)



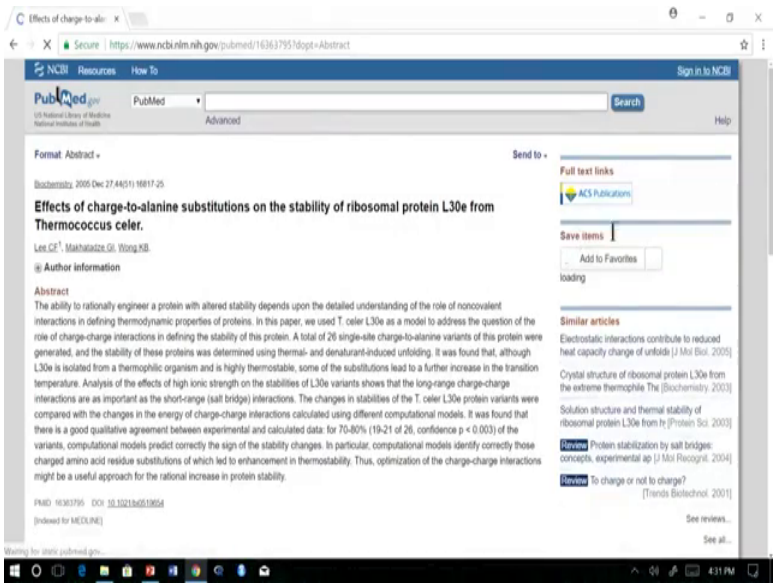
So, you can see this right. So, you can see the, this is the wild type and this is the mutation and, you can see this one.

(Refer Slide Time: 17:50)



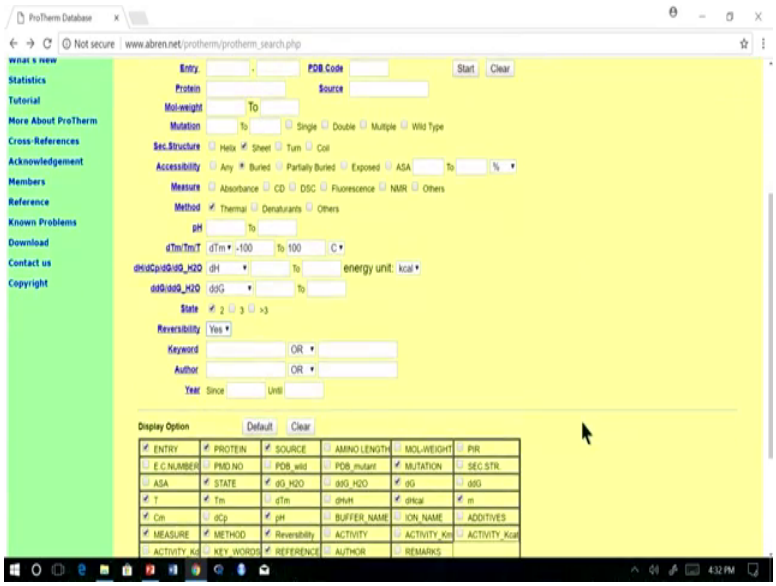
So, if you want to get the complete information you will go to the purple right. So, in the bio chemistry paper, if you see the purple you can access the abstract.

(Refer Slide Time: 18:03)



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)

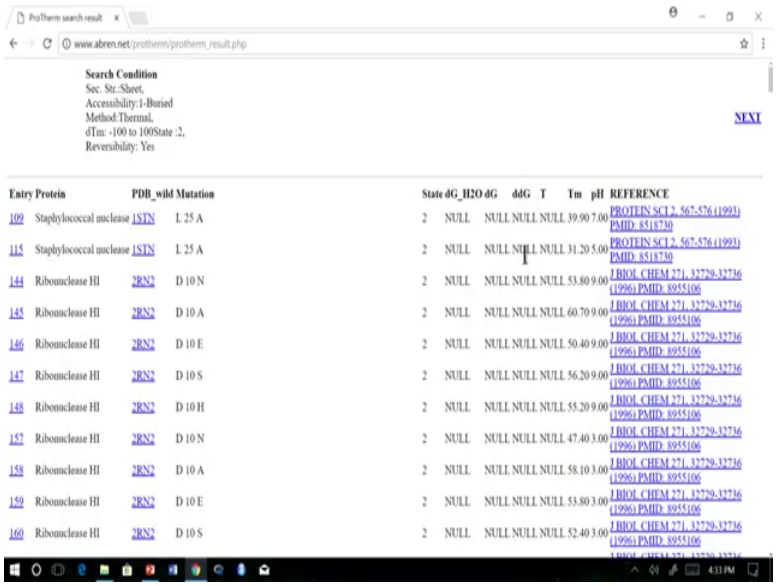


So, now if you want have multiple a choices for example, you are interested only on the buried mutations and in sheets and, obtained with the a thermal denaturation right, then you better to select the dTm, there is a range is minus 1 100 to plus 100 degree C right.

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.

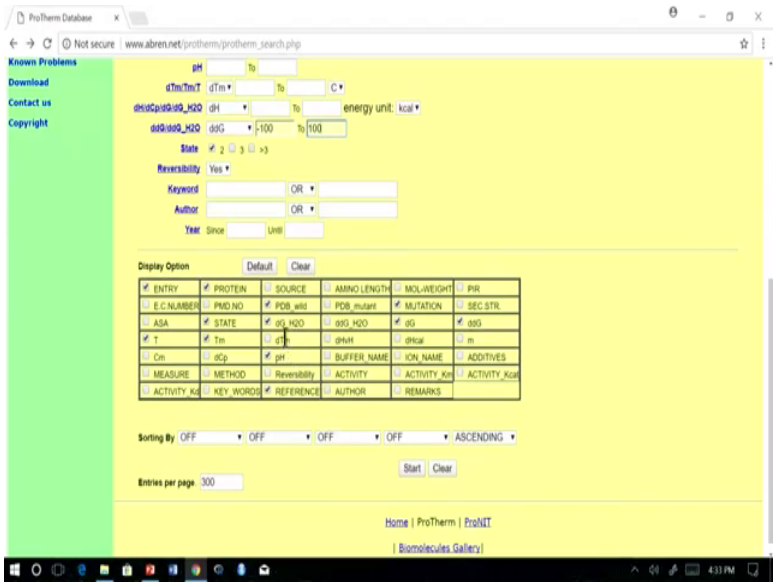
Then if you do like this and here we need to change because, we already it is a thermal it automatically it a put the Tm and the T is not Ta C Tm, we do not need and measure method, we reduce it reversibility is yes is the reference fine, there is m we do not need delta delta G, we need and this we remove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

(Refer Slide Time: 19:32)



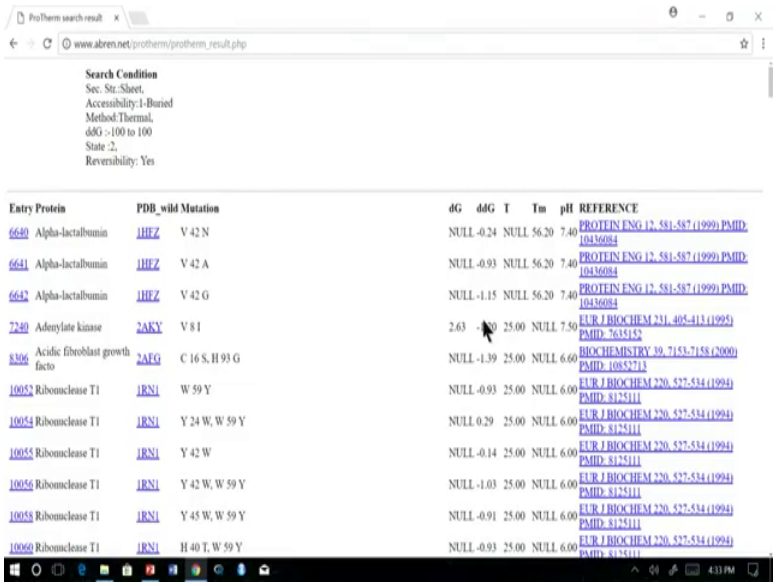
So, delta yes delta delta G did I use this yes is thermal.

(Refer Slide Time: 19:49)



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.

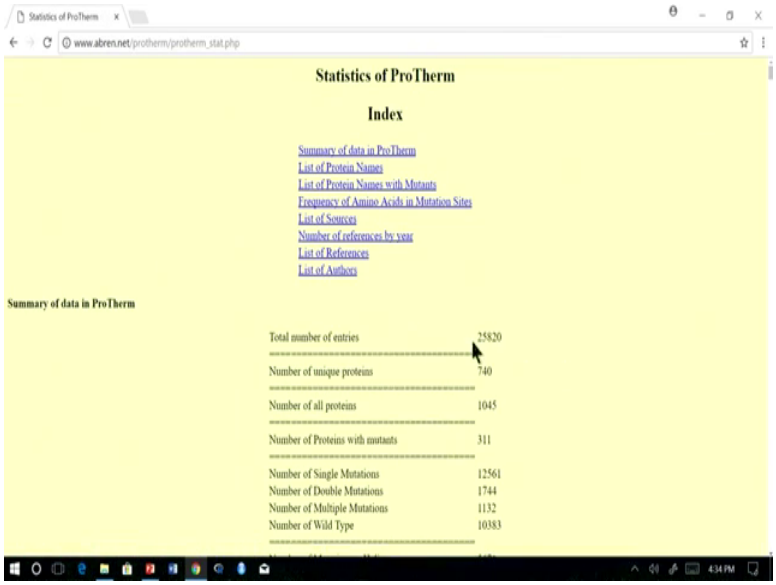
(Refer Slide Time: 20:16)



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.

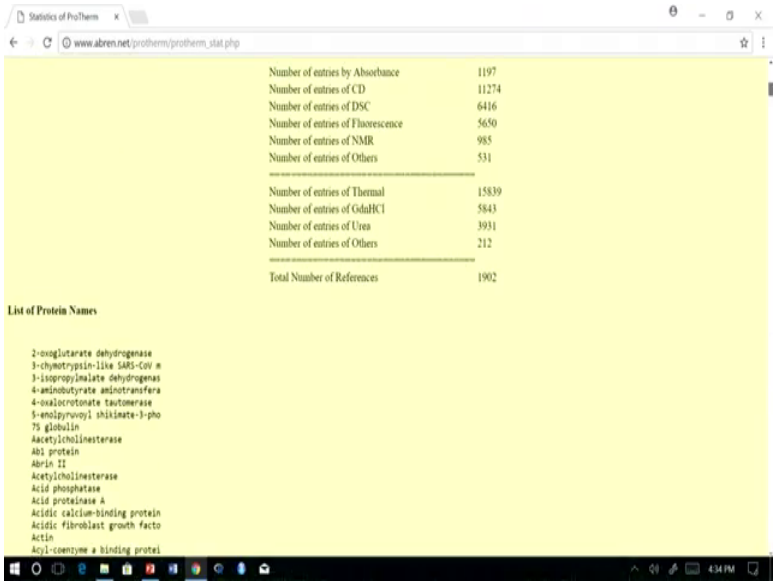
(Refer Slide Time: 20:31)



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

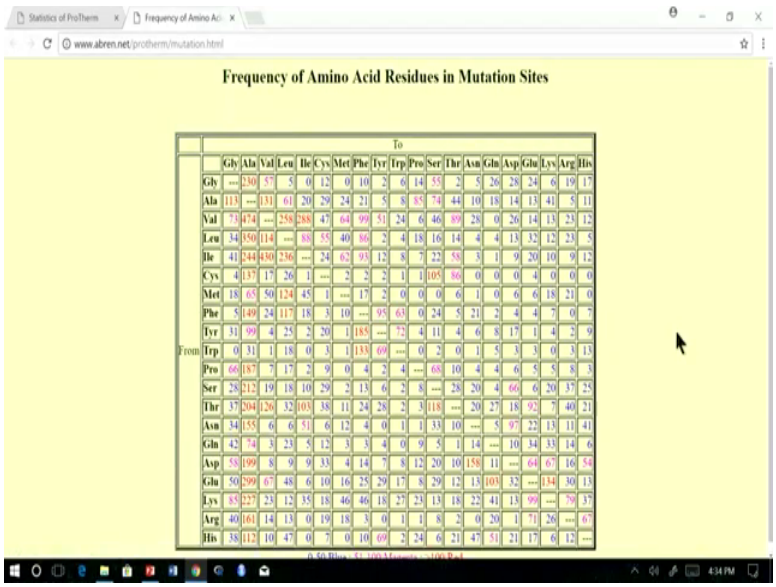
analyses. Let us say it is 740 unique proteins right then, they would 311 proteins content of the mutants.

(Refer Slide Time: 20:46)



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)

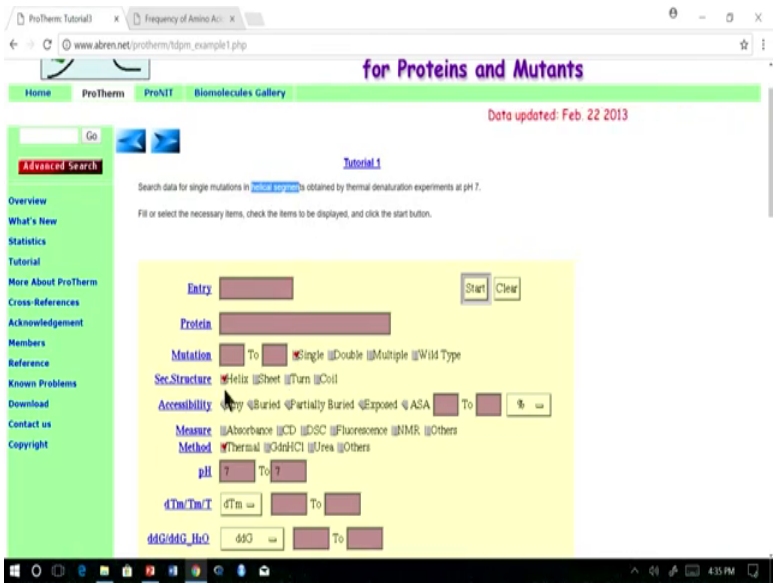


You can see the frequency table I showed this earlier so, you can see the some mutations are highly preferred because, they experimental lists they do not want to spend the

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.

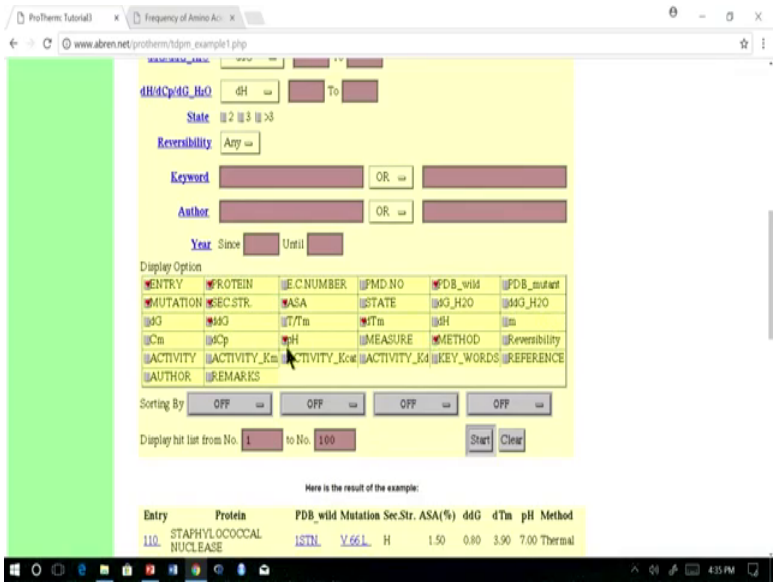
So, likewise you can get the statistics right. Now, if you go to the tutorial right. So, if it is a so, here we gave four different a questions first question is search a data for single mutation, in helical sequence obtained with thermal denaturation experiments, what will do with this what are the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

(Refer Slide Time: 21:49)



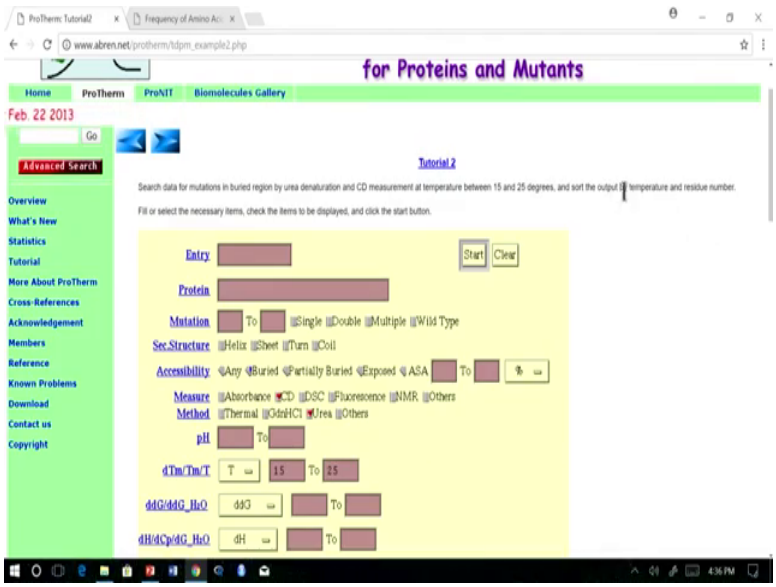
We need the because the question is helical sequence. So, we need to a click on a helical and the single mutation and, you can click on the single mutation right. And the accessibility does not matter because; we wrote we are not interested in we taking anything on the pH 7. So, you have got the pH 7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

(Refer Slide Time: 22:13)



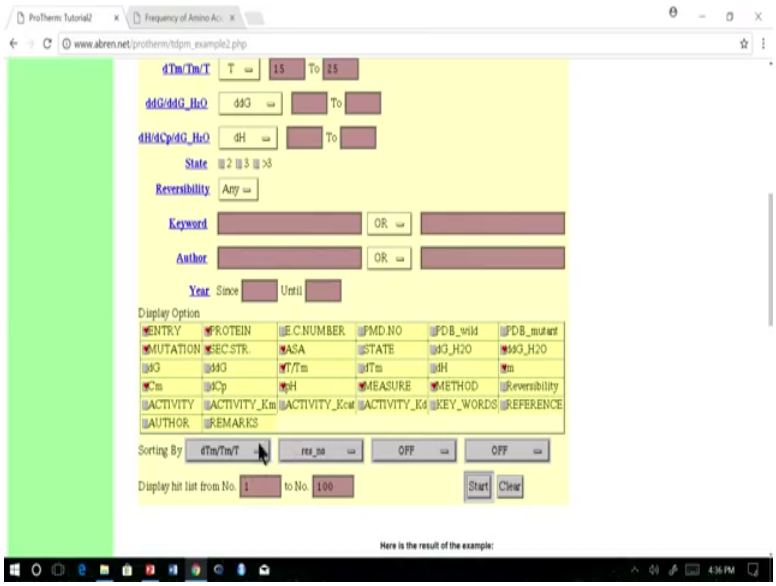
Right then, currently we click the display option, then we will start this then finally, we get the data now it is symbol the first one.

(Refer Slide Time: 22:24)



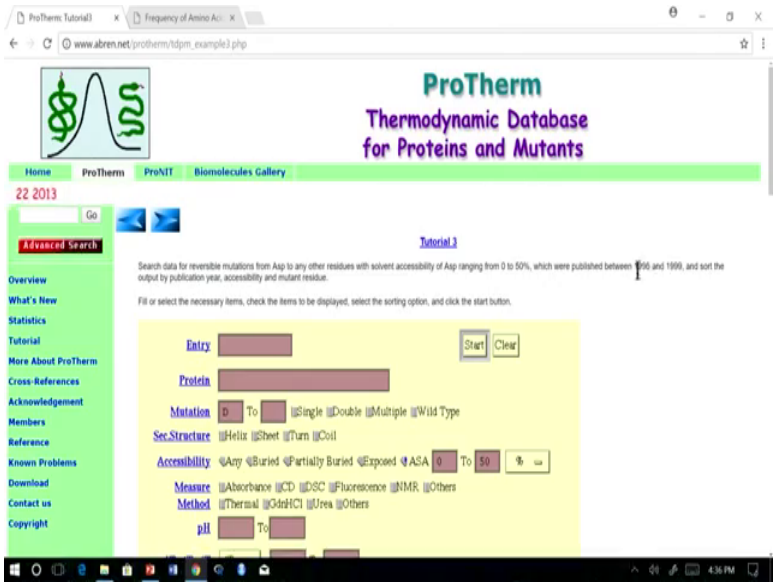
And go a second one this is with the denaturant denaturation here, we have the buried helix region, buried means, you have to get the accessibility buried right and, then you can see the urea denaturation. So, you put the urea and the CD measurement you put the CD and the temperature between 15 and 25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

(Refer Slide Time: 22:46)



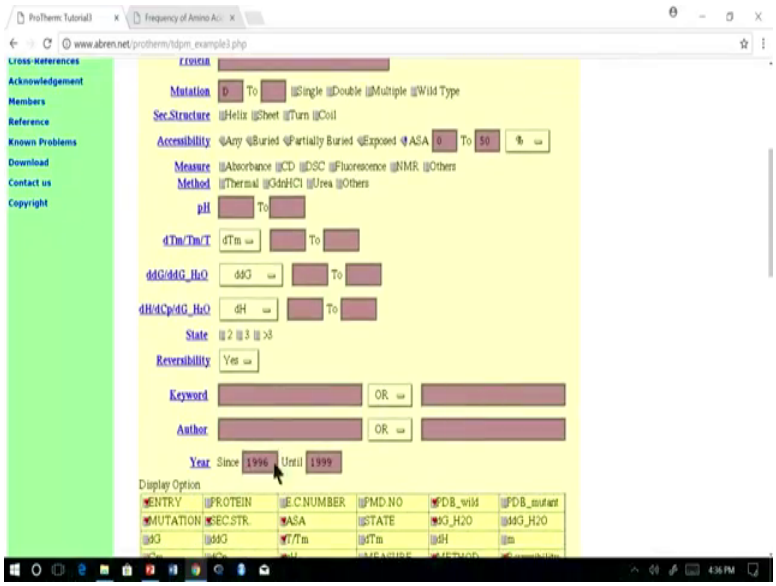
So, in this case sorting you need to change temperature residue number right and, if we start and we will get the data is the one for these things right. Then the will take one more example.

(Refer Slide Time: 23:01)



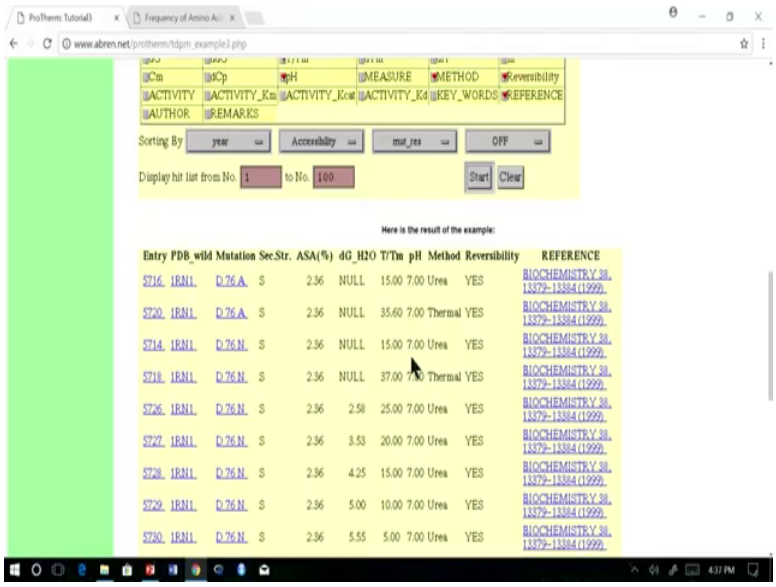
So, for example, reversible mutation and aspartic acid to any other residues; so, wild type residues aspartic acids now here is aspartic residues here, D to any other residue. So, in this case I put blank right, then you use the other conditions accessibility and the year of publication between 96 to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

(Refer Slide Time: 23:21)



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

(Refer Slide Time: 23:25)



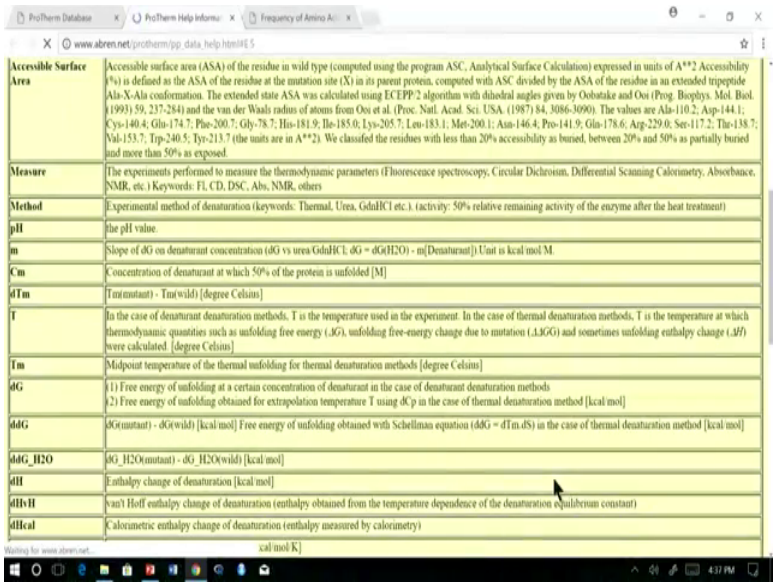
So, you have various options in ProTherm and you can use the multiple combination of different options to get your information fine. So, we use various terms right so, what is the meaning of accessibility and for example, I will put the buried and partially buried and exposed, what is the meaning of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



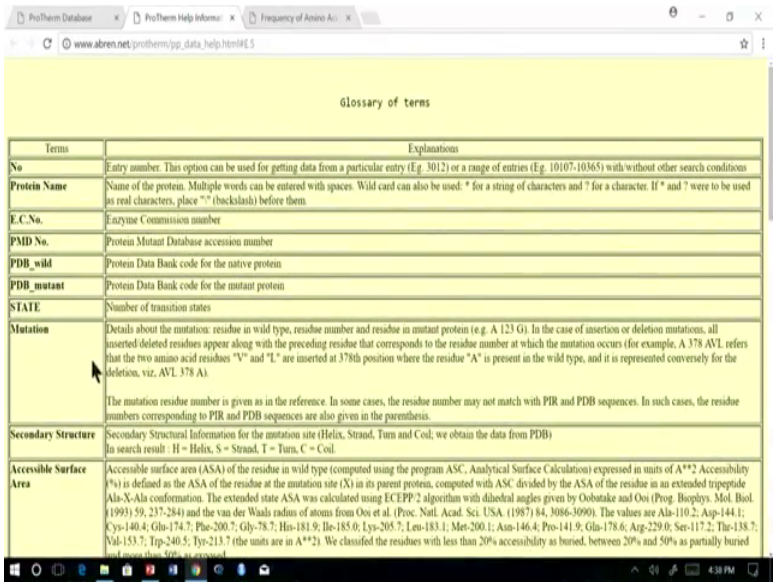
This will tell you what are the conditions, we used for this ProTherm, because we can use different cutoff values for the buried or partially exposed. In this case we use less than 20 percent buried that is very wide, and 20 to 50 as partially buried and more than will be exposed.

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)

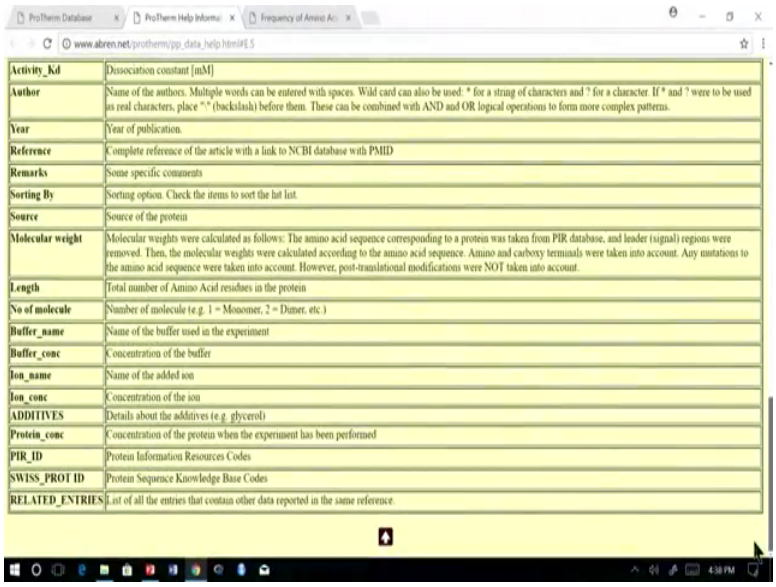


Glossary of terms

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 A**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 A**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)

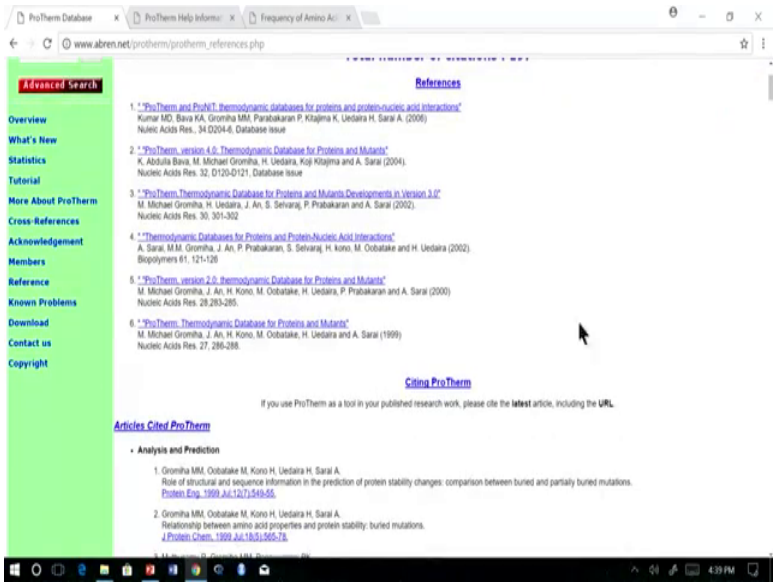


Activity_Kd	Dissociation constant [mM]
Author	Name of the authors. Multiple words can be entered with spaces. Wild card can also be used: * for a string of characters and ? for a character. If * and ? were to be used as real characters, place "\" (backslash) before them. These can be combined with AND 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 hit 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.

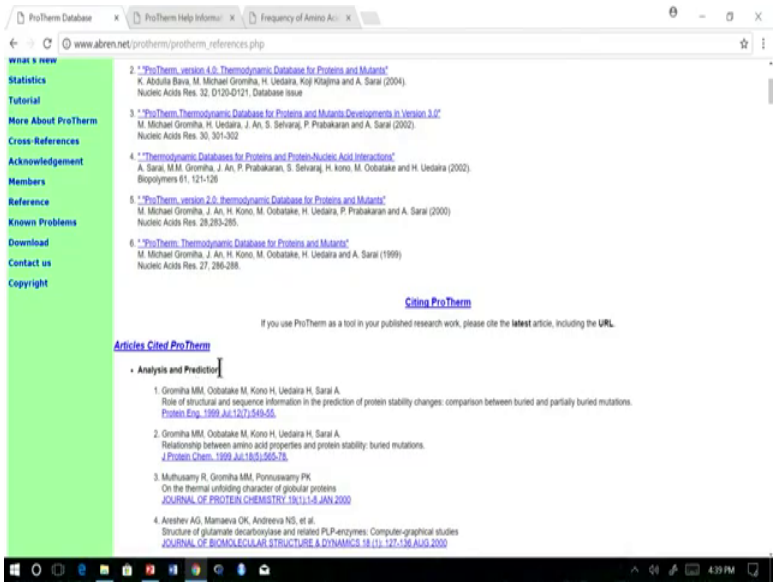
So, what are the applications of ProTherm. So, for what we use ProTherm there is no recent updates but if you go to the references ok, reference if you see.

(Refer Slide Time: 25:15)



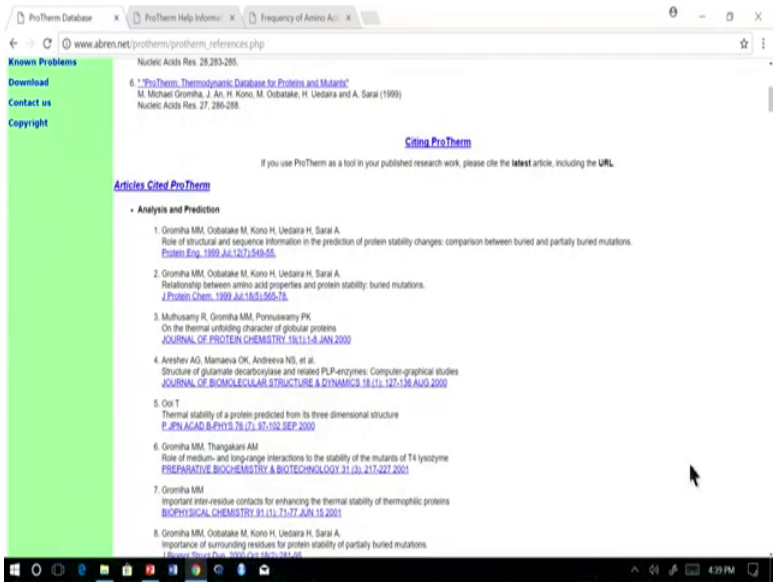
So, the various application these are the database papers.

(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.

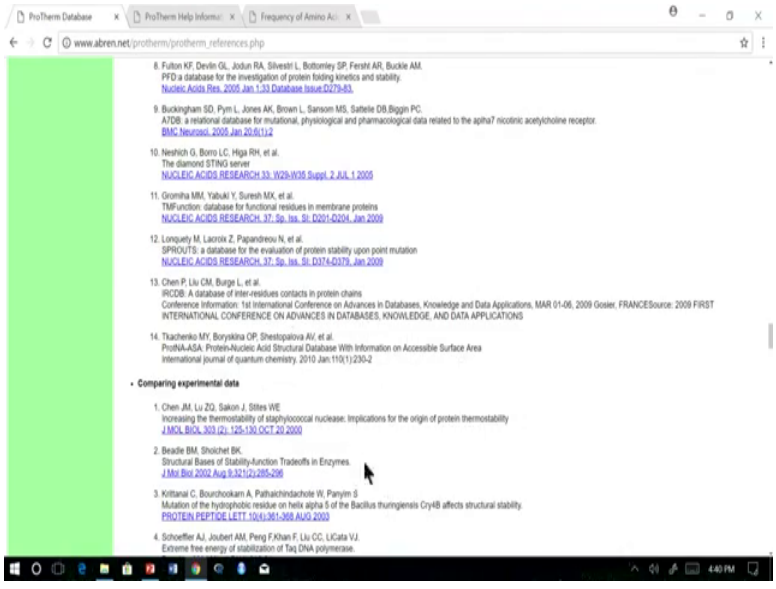
(Refer Slide Time: 26:06)

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.

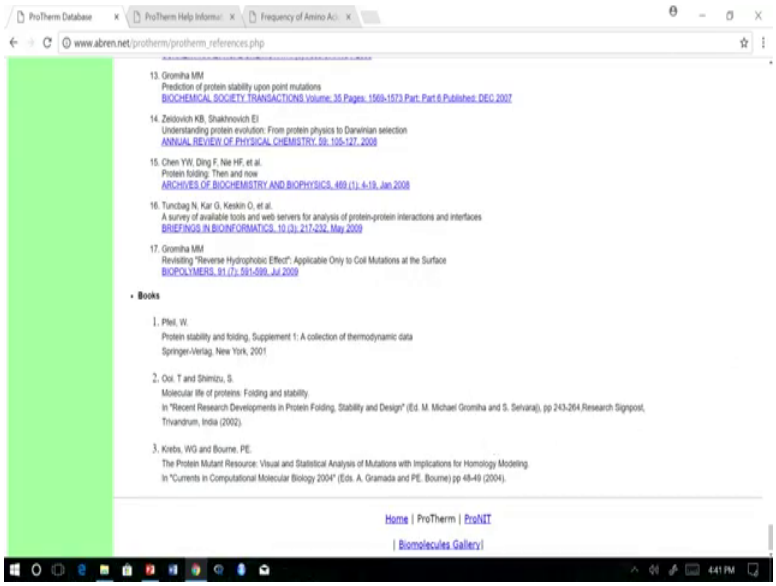
(Refer Slide Time: 26:45)



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 mutations in the ProTherm database or not right this is that, they use.

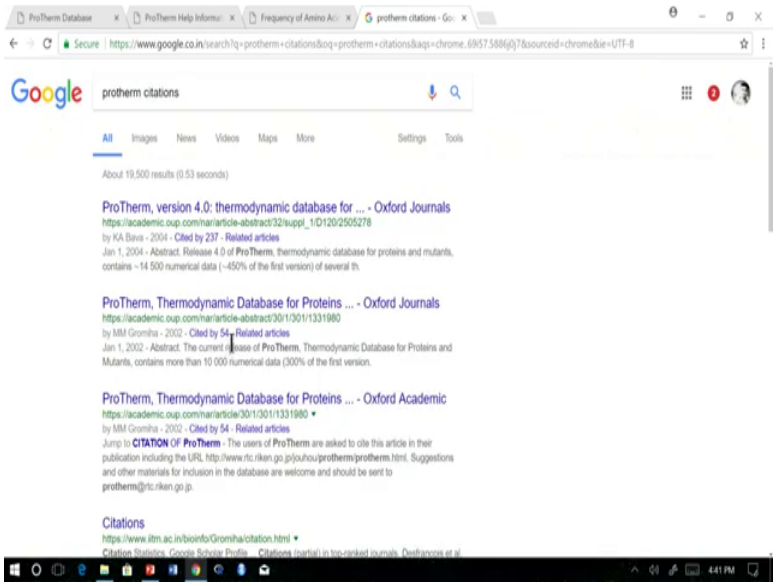
Now, there are many many citations, then also cited and very general applications and why are we can receive for the by technological applications, with increase in stability. So, right they are used the ProTherm database. So, there is several review article also they all sent the books.

(Refer Slide Time: 27:21)



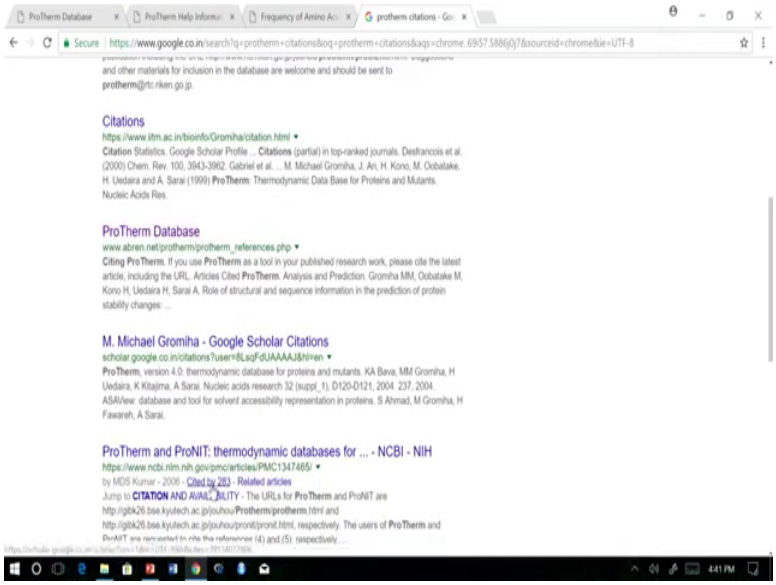
So, they also cite the ProTherm. So, this will tell the how the ProTherm is a important in this case.

(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)



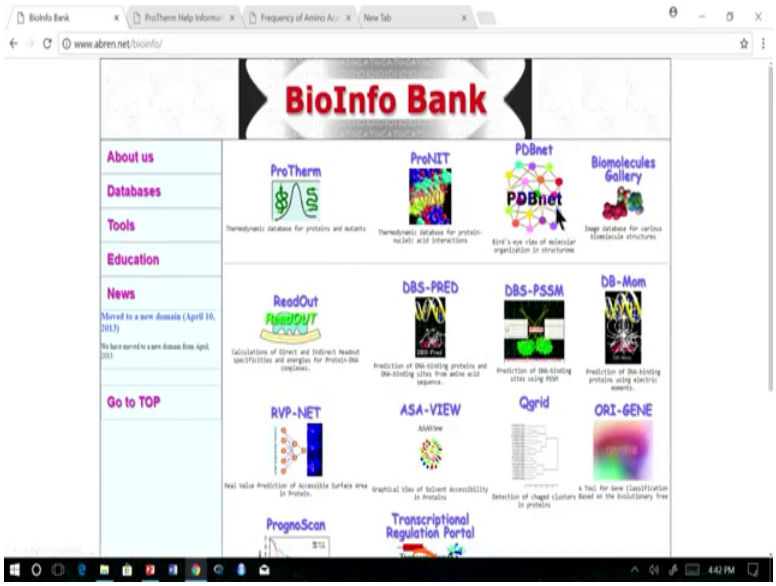
This is 283 citations to some citations, it totally in a if you add up all these 4 5 papers right.

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.