



Modifications are a very important topic in database searching.

In some cases, the main focus of a study is to characterise post translational modifications, which may have biological significance. Phosphorylation would be a good example.

In other cases, the modification may not be of interest in itself, but you need to allow for it in order to get a match. Oxidation during sample preparation would be an example.

And, of course, many methods of quantitation involve modifications containing isotopic labels

Some sequence variants, such as the substitution of one residue by another, are equivalent to modifications, and can be handled in a similar way

Addres	ss 🙋	http://v	ww.un	imod.o	rg/modifications_li	st.php?goto=14		- ja ja			~	ə G
	l	J٢	11)	0	DD prot	ein modification	s for mass spectro	ometry			Help	
		A	dd ew	oa s	earch for: [Logged a Any field	Contains	Change password Advanced search	arch Details S Page 1	found: 31 4 of 27	Records Per Page:: 20 V	
							Sele	ct/Unselect all Delete selected				
	ľ	•	, E	Î	Accession #	PSI-MS Name	Interim name	Description	Monoisotopic mass	Average mass	Composition	
	Edit	Сору	View		40	Sulfo	Sulfation	O-Sulfonation	79.956815	80.0632	O(3) S	
	Edit	Сору	View		21	Phospho	Phospho	Phosphorylation	79.966331	79.9799	H O(3) P	
	Edit	Сору	View		549		Cys->Trp	Cys->Trp substitution	83.070128	83.0670	H(5) C(8) N S(-1)	
	Edit	Сору	View		211	NEIAA	NEIAA-d0	N-ethyl iodoacetamide-d0	85.052764	85.1045	H(7) C(4) N O	
	Edit	Сору	View		747		Malonyl	Malonylation of C and S residues	86.000394	86.0462	H(2) C(3) O(3)	
	Edit	Сору	View		371	нмук	HMVK86	Michael addition of hydroxymethylvinyl ketone to cysteine	86.036779	86.0892	H(6) C(4) O(2)	
	Edit	Сору	View		324	DTBP	DTBP	dimethyl 3,3\'-dithiobispropionimidate	87.014270	87.1435	H(5) C(3) N S	
	Edit	Сору	View		178	DAET	ser_thr_DAET	phosphorylation to amine thiol	87.050655	87.1866	H(9) C(4) N O(-1) S	
	Edit	Сору	View		379	Hypusine	hypusine	hypusine	87.068414	87.1204	H(9) C(4) N O	
	Edit	Сору	View		126	Thioacyl	DSP	thioacylation of primary amines (N-term and Lys)	87.998285	88.1283	H(4) C(3) O S	
	Edit	Сору	View		185	Label:13C(9) +Phospho	13C9_Phospho_Tyr	C13 label (Phosphotyrosine)	88.996524	88.9138	H C(-9) 13C(9) O (3) P	
	Edit	Сору	View		212	NEIAA:2H(5)	NEIAA-d5	N-ethyl iodoacetamide-d5	90.084148	90.1353	H(2) 2H(5) C(4) N O	
	Edit	Сору	View		724		O-Methylphosphate	O-Methylphosphorylation	93.981981	94.0065	H(3) C O(3) P	
Done	8						<u> </u>				Internet	

Comprehensive and accurate information about post translational and chemical modifications is an essential factor in the success of protein identification. In Mascot, we take our list of modifications from Unimod, which is an on-line modifications database.

UNIM	OD	protein modifi	instigne for mass enastrom	ata,			-
Unimod, View I	record [A	ccession #: 5	6]	o u y		Help	
Back to list							
Accession #	56		PSI-MS Name	Acetyl:2H(3)	Interim Name	Acetyl_heavy	
Description	Acetate la	beling reagent (N	I-term & K) (heavy form, +3amu	(د			
Alt. Description	N-trideute	riumaceto×y					
Composition	H(-1) 2H(3	3) C(2) O	Monoisotopic	45.029395	Average	45.0552	
Specificity Def	inition 1						
Site	к		Position	Anywhere	Classification	Isotopic label	
Hidden	1		Group	1			
Specificity Def	inition 2				et	and the second second	
Site	N-term		Position	Any N-term	Classification	Isotopic label	
Notos and Pofe	1 Droncos		Group	2			
Source	PubMed PMID	Reference	11857757				
Source	PubMed PMID	Reference	11999733				
Source	PubMed PMID	Reference	12175151				
Source	Journal	Reference	Fred E., Department of Chemis	effects in comparative proteon stry, Purdue University, West L	nics. Zhang, Roujian; Sioma, Cathy afayette, IN, USA. Analytical Chemi:	S.; Thompson, Robert A.; Xiong, Li; Reg stry (2	hier,
Source	Journal	Reference	Global internal standard techno University, West Lafayette, IN,	ology for comparative proteom USA. Journal of Chromatogray	ics. Chakraborty, Asish; Regnier, Fr ohy, A (2002), 949(1-2), 173-184.	ed E Department of Chemistry, Purdue	
Source	Journal	Reference	Comparative proteomics based Peiran; Chakraborty, Asish; Se	d on stable isotope labeling and eeley, Erin; Sioma, Cathy; Tho	l affinity selection. Regnier, Fred E.; mpson, Robert A. Department of Ch	Riggs, Larry; Zhang, Roujian; Xiong, Li; emistry, Pu	Liu,
Curator	penner	Last Modified	2006-10-16 10:02:50		Verified	Yes	
Back to list							
							~
						Internet	

There are other lists of modifications on the web, like DeltaMass on the ABRF web site and RESID from the EBI, but none is as comprehensive as Unimod

Mass values are calculated from empirical chemical formulae, eliminating the most common source of error. Specificities can be defined in ways that are useful in database searching, and there is the option to enter mass-spec specific data, such as neutral loss information. This screen shot shows one of the better annotated entries, I can't pretend that all of them are this detailed. Nevertheless, it is a very useful, public domain resource that beats having to create your own list in an Excel spreadsheet or on the back of an envelope.



If you go to the help page, there is a link to download the contents of Unimod as a Mascot modifications file. This is the easiest way to keep the modifications list on an in-house Mascot server up-to-date



Here is a tip. The default list of modifications displayed in the Mascot search form is a short list, containing only the most common mods. If you want to see the complete list of mods, and you are using Mascot 2.2 or earlier, you need to follow the link at the bottom of the search form selection page

	Reader And Land And Land	🛛 🖾 🖉	2 🗾 🛄	PC	owermarks 🕅 🧥	*
idress 🙋 http://www.matri	xscience.com/cgi/form_defaults.pl					<u> </u>
(MATRIX) SCIENCE)		HOME WHAT'S	NEW I MASCOT I HELP I	PRODUCTSSUPP		Search Go
Mascot > Set Se	arch Defaults					
Set Mascot	search form defaults					
Database	MSDB					
Taxonomy	All entries	~				
Enzyme	Trypsin 💌					
Allow up to	1 v missed cleavages					
Fixed modifications	AB_old_ICATd0 (C) AB_old_ICATd8 (C) Acetyl (K) Acetyl (N-term) Amide (C-term)					
Variable modifications	AB_old_ICATd0 (C) AB_old_ICATd8 (C) Acetyl (K) Acetyl (N-term) Amide (c-term)					
Show all mods.						
ICAT	□ (MS/MS only)					
Peptide tol. ±	1.2 Da 💌					
MS/MS tol. ±	0.6 Da 💌					
Peptide charge	1+ 💌					
Monoisotopic						
Data format	Mascot generic 🛛 MS/MS only)					
						Internet

Check the box for Show all mods, then choose Save. This still sets the default state of the checkbox in Mascot 2.3, but we decided to place the checkbox on the search form, so as to make it easier to swap between the short and long lists.



It is extremely important that you do not choose more than the absolute minimum number of variable modification in a search. We talked about this in an earlier presentation, but it is worth repeating.

Variable or differential or non-quantitative modifications are expensive, in the sense that they increase the time taken for a search and reduce its specificity. This is because the software has to permute out all the possible arrangements of modified and unmodified residues that fit to the peptide molecular mass. As more and more modifications are considered, the number of combinations and permutations increases geometrically. The socalled combinatorial explosion.

Some variable modifications are worse than others. Modifications that only apply to a terminus, especially if they only apply when particular residue is at the terminus, like pyroglu, make little difference to the number of peptides to be tested. The problem modifications are the ones that apply to residues in any position, especially if they apply to multiple residues, like phosphorylation.

Unless you have enriched the sample in a particular PT-mod, e.g IMAC for phosphopeptides, it is usually not a good idea to try and catch PT-mods in a first pass search. Better to use a second pass search, which we call an error tolerant search, to catch the low abundance mods. We will come back to this later.



To illustrate this point. This search of a single MS/MS spectrum, using one variable mod, gives a nice, statistically significant match.

If the search is repeated with 8 mods, the match is the same, with an identical score, but now it is barely significant.

All of these mods have effectively increased the size of the database by a factor of 30

What's worse, the search takes over 10 times as long!

So, use variable mods sparingly. You'll get better results and faster.

By the way, the yellow region in the histogram indicates scores above the homology but below the identity thresholds. You will only see these regions highlighted in an MS/MS search report if it is a search of a single spectrum.



Of all post-translational modifications, phosphorylation is one of the most interesting and also one of the most difficult. Why is it such a challenge?

Peptide Su	unmary Report
Format As	Peptide Summary Help
	Significance threshold p< 0.05 Max. number of hits AUTO
	Standard scoring MudPIT scoring Ions score or expect cut-off Show sub-sets
	Show pop-ups ③ Suppress pop-ups ③ Sort unassigned Decreasing Score 🖌 Require bold red 🗌
(a)	
Select All	Select None Search Selected Liror toler ant
1. CASE	BOVIN Mass: 25091 Score: 88 Matches: 1(1) Sequences: 1(1)
Beta	-casein OS=Bos taurus GN=CSN2 PE=1 SV=2
Beta Checi	-casein OS-Bos taurus GN=CSN2 PE=1 SV=2 k to include this hit in error tolerant search
Beta	Casein OS-Bos taurus GN=CSN2 PE-1 SV=2 k to include this hit in error tolerant search
Beta Check Query	casein OS=Bos taurus GN=CSN2 PE=1 SV=2 k to include this hit in error tolerant search 7 Observed Mr(expt) Mr(calc) Delta Miss Score Expect Rank Unique Peptide 1 13114000 2660.7854 2060.8212 -0.0337 0 88 1.6e-06 1 V K.FOSEE000TEDELODK.I + Phose
Beta Check Query	-camein OS-Bos taurus GN=CSN2 PE=1 SV=2 k to include this hit in error tolerant search y Observed Mr(expt) Mr(calc) Delta Miss Score Expect Rank Unique Peptide 1031.4000 2060.7854 2060.8212 -0.0357 0 88 1.6e-06 1 V K.FQSEEQQ0TEDELQDK.I + Phosy
Beta Check Query	-casein OS-Bos taurus GN=CSN2 PE=1 SV=2 k to include this hit in error tolerant search y Observed Mr(expt) Mr(calc) Delta Miss Score Expect Rank Unique Peptide 1031.4000 2060.7854 2060.8212 -0.0357 0 88 1.6e-06 1 U K.FQSEEQQQTEDELQDK.I + Phosp
Beta Check Query V Prot	casein OS-Bos taurus GN=CSN2 PE-1 SV=2 k to include this hit in error tolerant search v Observed Mr(expt) Mr(calc) Delta Miss Score Expect Rank Unique Peptide 1031.4000 2060.7854 2060.8212 -0.0357 0 88 1.6e-06 1 U K.FQSEEQQQTEDELQDK.I + Phosp Top scoring peptide matches to query 1 Score greater than 39 indicates homology
Beta Checi Query ✓ Prot	-casein OS-Bos taurus GN=CSN2 PE-1 SV=2 k to include this hit in error tolerant search y Observed Mr(expt) Mr(calc) Delta Miss Score Expect Rank Unique Peptide 1031.4000 2060.7854 2060.8212 -0.0357 0 88 1.6e-06 1 V K.FQSEEQQQTEDELQDK.I + Phosp Top scoring peptide matches to query 1 Score greater than 39 indicates homology Score greater than 39 indicates dentity
Beta Checi Query ✓ Prote Beta	-casein 05=Bos taurus GN=CSN2 PE=1 SV=2 At to include this hit in error tolerant search 7 Observed Mr(expt) Mr(calc) Delta Miss Score Expect Rank Unique Peptide 1031.4000 2060.7854 2060.8212 -0.0357 0 88 1.6e-06 1 U K.FQSEEQQQTEDELQDK.I + Phosp Top scoring peptide matches to query 1 Score greater than 39 indicates homology Score greater than 43 indicates identity France France France Content of the State of t
Beta Checi Query V Prote <u>CASB</u> Beta CASB	-casein OS-Bos taurus GN-CSN2 PE-1 SV-2 k to include this hit in error tolerant search 9 Observed Mr(expt) Mr(calc) Delta Miss Score Expect Rank Unique Peptide 1031.4000 2060.7854 2060.8212 -0.0357 0 88 1.6e-06 1 U K.FQSEE000TEDELQDK.I + Phosp Top scoring peptide matches to query 1 Score greater than 39 indicates identity Score greater than 43 indicates identity Score Expect Delta Mit Protein Peptide 88.3 1.6e-06 -0.0357 1 C.SR BOUNK K.FOGEFOOGTEDELODE, I
Beta Checi Query Prote CASB Beta CASB	-casein OS-Bos taurus GN=CSN2 PE-1 SV=2 k to include this hit in error tolerant search y Observed Mr(expt) Mr(calc) Delta Miss Score Expect Rank Unique Peptide 1031.4000 2060.7854 2060.8212 -0.0357 0 88 1.6e-06 1 U K.FQSEEQQQTEDELQDK.I + Phosp Top scoring peptide matches to query 1 Score greater than 39 indicates homology Score greater than 43 indicates identity Score Expect Delta Mit Protein Peptide 88.3 1.6e-06 -0.0357 1 CASE BOVIN K.FQSEEQQQTEDELQDK.I 28.6 1.5 -0.0357 1 CASE BOVIN K.FOGEEQQCTEDELQDK.I 28.6 1.5 -0.0357 1 CASE DOVIN K.FOGEEDQCTEDELQDK.I
Beta Checi Query V Prot CASB Beta Beta Beta	-casein 05=Bos taurus GN=CSN2 PE-1 SV=2 k to include this hit in error tolerant search v Observed Mr(expt) Mr(calc) Delta Miss Score Expect Rank Unique Peptide 1031.4000 2060.7854 2060.8212 -0.0357 0 88 1.6e-06 1 V K.FQSEEQQQTEDELQDK.I + Phosp Top scoring peptide matches to query 1 Score greater than 39 indicates identity Score greater than 43 indicates identity Score Expect Delta Hit Protein Peptide 88.3 1.6e-06 -0.0357 1 CASE_BOVIN K.FQSEEQQQTEDELQDK.I 28.6 1.5 -0.0357 1 CASE_BOVIN K.FQSEEQQQTEDELQDK.I 21.0 8.8 -0.1866 K.CLSLSKYOULPFETIEK.H
Beta Check Query Prote Beta CASB Beta	-casein OS-Bos taurus GN-CSN2 PE-1 SV-2 k to include this hit in error tolerant search y Observed Mr(expt) Mr(calc) Delta Miss Score Expect Rank Unique Peptide 1031.4000 2060.7854 2060.8212 -0.0357 0 88 1.6e-06 1 V K.PQSEEQQ0TEDELQDK.I + Phosp Top scoring peptide matches to query 1 Score greater than 39 indicates homology Score greater than 39 indicates identity Score Expect Delta Hit Protein Peptide 88.3 1.6e-06 -0.0357 1 CASE BOVIN K.PQSEEQQ0TEDELQDK.I 28.6 1.5 -0.0357 1 CASE BOVIN K.PQSEEQQTEDELQDK.I 28.6 1.5 -0.0357 1 CASE BOVIN K.PQSEEQQTEDELQDK.I 21.0 8.8 -0.1886 K.CLSLSKQPDLFETIEK.H 15.9 28 -0.0907 K.CUVDKSPMPEPEDEK.G
Beta Checi Query V Prot. CASB Beta CASB Beta	-casein OS-Bos taurus GN=CSN2 PE-1 SV=2 At to include this hit in error tolerant search V Observed Mr(expt) Mr(calc) Delta Miss Score Expect Rank Unique Peptide 1031.4000 2060.7854 2060.8212 -0.0357 0 88 1.6e-06 1 U K.FQSEEQQQTEDELQDK.I + Phosp Top scoring peptide matches to query 1 Score greater than 39 indicates homology Score greater than 39 indicates identity Score Expect Delta Mit Protein Peptide 88.3 1.6e-06 -0.0357 1 CASE BOVIN K.FQSEEQQQTEDELQDK.I 28.6 1.5 -0.0357 1 CASE BOVIN K.FQSEEQQQTEDELQDK.I 21.0 0.8 -0.1886 K.CLSLSRQUETETIEK.H 15.9 28 -0.0907 K.QMVDKDSPHVEPDEK.G 14.1 42 -0.1713 K.QLASCFTHOREQKQAK.R
Beta Check Query Prot. CASB Beta CASB Beta Search Par	-casein OS-Bos taurus GN-CSN2 PE-1 SV-2 k to include this hit in error tolerant search y Observed Mr(expt) Mr(calc) Delta Miss Score Expect Rank Unique Peptide 1031.4000 2060.7854 2060.8212 -0.0357 0 80 1.6e-06 1 U K.FQSEEQQ0TEDELQDK.I + Phosp Top scoring peptide matches to query 1 Score greater than 39 indicates inducing Score greater than 39 indicates inducing Score expect Delta Hit Protein Peptide 88.3 1.6e-06 -0.0357 1 CASB_BOVIN K.FQSEEQQ0TEDELQDK.I 28.6 1.5 -0.0357 1 CASB_BOVIN K.FQSEEQQ0TEDELQDK.I 21.0 0.8 -0.1886 K.CLSLSKQVDIFETIEK.H 15.9 28 -0.0907 K.GMVVDKOBPHOFENC.G 14.1 42 -0.1713 K.GLASOFFINGEONKR.R
Beta Checi Query Prote CASB Beta CASB Beta Search Pai	-casein 05=Bos taurus GN=CSN2 PE-1 5V=2 & to include this hit in error tolerant search 7 Observed Mr(expt) Mr(calc) Delta Miss Score Expect Rank Unique Peptide 1031.4000 2060.7854 2060.8212 -0.0357 0 88 1.6e-06 1 V K.FQSEEQQQTEDELQDK.I + Phosp Top scoring peptide matches to query 1 Score greater than 39 indicates homology Score greater than 43 indicates identity Score Expect Delta Mit Protein Peptide 88.3 1.6e-06 -0.0357 1 CASE BOVIN K.FQSEEQQQTEDELQDK.I 28.6 1.5 -0.0357 1 CASE BOVIN K.FOSEEQQQTEDELQDK.I 28.6 1.5 -0.0357 1 CASE BOVIN K.FOSEEQQTEDELQDK.I 28.6 1.5 -0.1866 K.CUSLSKOWDLFEFIEK.H 15.9 28 -0.0907 K.CWNWDKOSPHEFEDEK.G 14.1 42 -0.1713 K.CULAGCETEINQEQKOAK.R 13.6 47 -0.1469 K.ITFLELTFRQONKLS.S 12.8 58 -0.1366 K.SSOUPTOPYTERFYK.6
Beta Check Query Frot. CASB Beta CASB Beta Search Par Type of sea	-casein OS-Bos taurus GN-CSN2 PE-1 SV-2 k to include this hit in error tolerant search y Observed Mr(expt) Mr(calc) Delta Miss Score Expect Rank Unique Peptide 1031.4000 2060.7854 2060.8212 -0.0357 0 88 1.6e-06 1 U K.FQSEEQQQTEDELQDK.I + Phosp Top scoring peptide matches to query 1 Score greater than 39 indicates identity Score greater than 43 indicates identity Score sequent Delta Mit Protein Peptide 88.3 1.6e-06 -0.0357 1 CASE BOVIN K.FQSEEQQQTEDELQDK.I 28.6 1.5 -0.0357 1 CASE BOVIN K.FQSEEQQQTEDELQDK.I 21.0 6.8 -0.1866 K.CSLSSEQUETELDK.E 14.1 42 -0.1713 K.CLSLSSEQUETELDK.E 15.6 47 -0.1469 K.ITPLEEITENQERQAK.R 15.8 58 -0.1366 K.SSSQLFTQPPTREPTOK.G 15.8 54 -0.2007 K.SLQCECOLVALDELAR.R
Beta Check Query Prot. CASB Beta Search Pal Type of sea Enzyme	-cazein OS-Bos taurus GN-CSN2 PE-1 SV-2 k to include this hit in error tolerant search y Observed Mr(expt) Mr(calc) Delta Miss Score Expect Rank Unique Peptide 1031.4000 2060.7854 2060.8212 -0.0357 0 88 1.6e-06 1 V K.PQSEEQQQTEDELQDK.I + Phosp Top scoring peptide matches to query 1 Score greater than 39 indicates homology Score greater than 39 indicates identity Score Expect Delta Hit Protein Peptide 88.3 1.6e-06 -0.0357 1 CASE BOVIN K.PQSEEQQQTEDELQDK.I 28.6 1.5 -0.0357 1 CASE BOVIN K.PQSEEQQTEDELQDK.I 28.6 1.5 -0.0357 1 CASE BOVIN K.PQSEEQQTEDELQDK.I 29.8 -0.1886 K.CLSLSKQVDLFETIEK.H 15.9 28 -0.0367 K.CMVVKDSPHEFEDEK.G 14.1 42 -0.1713 K.CLASEVTINDEQKAK.R 13.6 47 -0.1469 K.ITPLEEIYFNEDDINEK.S 12.8 58 -0.1366 K.SSSOTTOPPTYREFFOR.G 12.3 64 -0.2007 K.SUCREGNEWAEDLESK.R 1.9 71 -0.1635 K.VULLCVGETINEEDEK.T 11.5 72 -0.055 P. ONINGENERIE
Beta Checi Query Prot CASB Beta CASB Beta Search Pat	-casein OS-Bos taurus GN-CSN2 PE-1 SV-2 k to include this hit in error tolerant search y Observed Mr(expt) Mr(calc) Delta Miss Score Expect Rank Unique Peptide 1031.4000 2060.7854 2060.8212 -0.0357 0 88 1.6e-06 1 U K.F0SEE000TEDELQDK.I + Phosp Top scoring peptide matches to query 1 Score greater than 39 indicates homology Score greater than 43 indicates identity Score greater than 43 indicates identity Score greater than 41 indicates identity Score greater than 43 indicates identity Score action 1 CASE BOVIN K.F0SEE000TEDELQDK.I 28.6 1.5 -0.0357 1 CASE BOVIN K.F0SEE000TEDELQDK.I 21.0 6.8 -0.0866 K.CLSLSKOULPETLIEK.H 15.9 28 -0.0907 K.CULSKKOULPETLIEK.H 13.6 47 -0.1469 K.ITFLELVFNDONNEK.S 12.6 58 -0.1366 K.SSUCPTOPYDE0FOK.G 12.3 64 -0.2007 K.SUCEFCEDLSVAEDELSEK.R 11.9 71 -0.1635 K.VLLCVCETUREPEDEK.R 11.9 71 -0.1635 K.OUTLEVERDONNER.SK.R 11.5 78 -0.0655 R.ODFLDILLSAKSENTK.D

Lets look at an example or two.

One of the most common phosphopeptides comes from the milk protein, beta casein. There are two potential phosphorylation sites, S and T, but only one is modified. Because the two sites are widely separated, the two arrangements get very different scores.



Beautiful spectrum; long run of y ions; move site to T9 and many matches would disappear



Mascot 2.4 reports site localisation probabilities using the delta score method published in MCP by Bernard Kuster's group. They analysed a collection of synthetic analogs of real phosphopeptides and determined what score difference was required to determine the correct site with an error rate of (say) 5%. Because we don't expect everyone to calibrate their data in this way, we have made the calculation slightly more conservative. A score difference of 10 would give approximately 90% probability that the higher scoring arrangement was correct.

00 Error (ppn) 0 0				······	
RMS err	or 60 ppm BLAST cear	ch of FC	SEECOOTEDELODK	Mass (Da)	
Paran Other	eters: blastp,	nr protei	in database, expect=20000, no i	filter, PAM30)	
Allma	tches to this	ouerv	E		
Score	Mr(calc)	Delta	Sequence	Site Analysis	
88.3	2060.8212	-0.0357	FOSEEQOOTEDELODK	Phospho S3 100.00%	
28.6	2060.8212	-0.0357	FOSEEOOOTEDELODK	Phospho T9 0.00%	
21.0	2060.9741	-0.1886	CLSLSKOVDLFEETIEK		R
15.9	2060.8762	-0.0907	OMVVDKDSPHVEPEDEK		
14.1	2060.9568	-0.1713	<u>OLASGEYFLNOEOKOAK</u>		
13.6	2060.9343	-0.1489	ITFLEELYPKDODNEK		
12.8	2060.9221	-0.1366	SSSQIPTOPPVTK SPYGK		
12.3	2060.9862	-0.2007	SLQEGEGDLSVAEDRLSEK		
11.0	2060.9489	-0.1635	YLILCVGETILNERDSEK		
11.2	2060.8509	-0.0655	<u>QDFLDILLSAK SENTK</u>		
11.5			24		
11.5			Masc	ot: http://www.matrixsci	ence.com/
11.5					

A very large score difference such as the one we were just looking at gives 100% likelihood that the phosphate is on S3.

	II () www.inaut.sc			-1,0000,20120701,100111	iow i i dat	M
Peptide Su	mnary Report					
Format As	Peptide Summary	~			Help	
	Significance threshold	p< 0.05	Max. number o	f hits 20		
	Standard scoring ()	MudDIT cooring		meet out off 0	Shorr enh-sets 0	
		Mudrii sconig			Show sub-sets 0	
	Show pop-ups 💿 Su	appress pop-ups	 Sort unassigned 	Decreasing Score	Require bold red 🛄	
Select All	Select None Sea	rch Selected	Firor tolerant			
L. <u>KAPC</u>	BOVIN Mass: 4	0594 Score	: 79 Matche	es: 1(1) Sequences:	1(1)	
CAMP-	dependent protein	h kinase catal	ytic subunit al	lpha OS=Bos taurus G	N=PRK&C& PE=1 SV=3	
Chec)	to include this	hit in error	tolerant search	h		
Query	Observed Mr(expt) Mr(ca	lc) Delta Mi	ss Score Expect Ra	nk Unique Peptide	
✓ 4	1107.9039 2213	.7933 2214.0	683 -0.2750	0 80 8.5e-06	U R. <u>T</u> WTLCGT	PEYLAPEIILSK.G + P
	4	ide matches to	o guery 1			
	Ton scoring nent:					
Prote	Top scoring pept: Score greater the	an 30 indicate	es homology			
Prote	Top scoring pept: Score greater the Score greater the	an 30 indicato an 42 indicato	es homology es identity			
Prote KAPCJ CAMP-	Top scoring pept: Score greater the Score greater the	an 30 indicato an 42 indicato	es homology es identity	Pentide	PREACA PE=2	SV=3
Prote KAPCA CAMP-	Top scoring pept: Score greater the Score greater the Score Expect 80.4 8.5e-06	an 30 indicat an 42 indicat Delta Hi -0.2750	es homology es identity t Protein 1 KAPCA BOVIN	Peptide R.TWTLCGTPEYLAPEIII	PRKACA PE=2	SV=3
Proto <u>KAPC</u> CAMP- <u>KAPC</u>	Top scoring pept: Score greater the Score greater the Score Expect 80.4 8.5e-06 76.9 1.9e-05	an 30 indicat an 42 indicat Delta Hi -0.2750 -0.2750	es homology es identity t Protein 1 KAPCA_BOVIN 1 KAPCA_BOVIN	Peptide R. <u>T</u> WTLCGTPEYLAPEIII R.TW <u>T</u> LCGTPEYLAPEIII	PRKACA PE=2	SV=3
Prote KAPC, CAMP- KAPC, CAMP-	Top scoring pept: Score greater the Score greater the Score Expect 80.4 8.5e-06 76.9 1.9e-05 38.7 0.13	an 30 indicate an 42 indicate -0.2750 -0.2750 -0.2750	es homology es identity t Protein 1 KAPCA_BOVIN 1 KAPCA_BOVIN 1 KAPCA_BOVIN	Peptide R. <u>T</u> WTLCGTPEYLAPEIII R.TW <u>T</u> LCGTPEYLAPEIII R.TWTLCG <u>T</u> PEYLAPEIII	PRKACA PE=2 .SK.G .SK.G N=PRKACA PE=	SV=3 2 SV=2
Ргоt с <u>КАРС</u> сАМР- <u>КАРС</u> сАМР- <u>КАРС</u>	Top scoring pept: Score greater the Score greater the 80.4 8.5e-06 76.9 1.9e-05 38.7 0.13 18.0 15	an 30 indicate an 42 indicate Delta Hi -0.2750 -0.2750 -0.2750 -0.2750	es homology es identity t Protein 1 KAPCA_BOVIN 1 KAPCA_BOVIN 1 KAPCA_BOVIN 1 KAPCA_BOVIN	Peptide R.TWTLCGTPEYLAPEIII R.TWTLCGTPEYLAPEIII R.TWTLCGTPEYLAPEIII R.TWTLCGTPEYLAPEII	PRKACA PE=2 .SK.G .SK.G N=PRKACA PE= .SK.G	SV=3 2 SV=2
Ргоt е <u>КАРС</u> , сАМР- <u>КАРС</u> , сАМР- сАМР-	Top scoring pept: Score greater the Score greater the Score Expect 80.4 8.5e-06 76.9 1.9e-05 38.7 0.13 18.0 15 12.6 51 12.6 51	an 30 indicate an 42 indicate Delta Hi -0.2750 -0.2750 -0.2750 -0.2750 -0.2750 -0.2750 -0.2111	es homology es identity t Protein 1 KAPCA_BOVIN 1 KAPCA_BOVIN 1 KAPCA_BOVIN 3 GSA_XYLFT 2 CSA_YVLFT	Peptide R.TWILCGTPEYLAPEIII R.TWILCGTPEYLAPEIII R.TWILCGTPEYLAPEIII K.GGSGMLTLGTPSSPGV P.GGSGMLTLGTPSSPGV	PRKACA PE=2 .SK.G .SK.G .SK.G N=PRKACA PE= .SK.G .AELSK.L CA PE=1 SV=2	SV=3 2 SV=2
Prote KAPC, CAMP- KAPC, CAMP- KAPC, KAPC,	Top scoring pept: Score greater that Score greater that Score Expect 80.4 8.5e-06 76.9 1.9e-05 38.7 0.13 18.0 15 12.6 51 12.6 51 12.6 51	an 30 indicat. an 42 indicat. Delta Him -0.2750 -0.2750 -0.2750 -0.2750 -0.2750 -0.2111 -0.2111 -0.2111	mology identity KAPCA_BOVIN KAPCA_BOVIN KAPCA_BOVIN KAPCA_BOVIN KAPCA_BOVIN GSA_XYLFT GSA_XYLFT GSA_YVLFT	Peptide R.TWILCGTPEYLAPEIII R.TWILCGTPEYLAPEIII R.TWILCGTPEYLAPEII R.TWILCGTPEYLAPEII K.GGSGHLTLGIPSSPOH K.GGSGHLTLGIPSSPOH K.GGSGHLTLGIPSSPOH	PRKACA PE=2 .SK.G .SK.G .SK.G .SK.G ALSK.L CA PE=1 SV=2 AELSK.L DETSV.L	SV=3 2 SV=2
Prote KAPC CAMP- KAPC CAMP- KAPC CAMP- KAPC CAMP-	Top scoring pept: Score greater th Score greater th 80.4 8.5e-06 76.9 1.9e-05 38.7 0.13 18.0 15 12.6 51 12.6 51 12.6 51 12.6 51	an 30 indicat. Delta Hi -0.2750 -0.2750 -0.2750 -0.2750 -0.2111 -0.2111 -0.2111 -0.2111	es homology es identity t Protein 1 KAPCA_BOVIN 1 KAPCA_BOVIN 1 KAPCA_BOVIN 3 GSA_XYLFT 3 GSA_XYLFT 2 GSA_XYLFT 2 GSA_XYLFN	Peptide R.TWILCGTPEYLAPEIII R.TWILCGTPEYLAPEII R.TWILCGTPEYLAPEII K.GGSGHLILGIPSSPOVI K.GGSGHLILGIPSSPOVI K.GGSGHLILLGIPSSPOVI	PRKACA PE=2 .SK.G .SK.G .SK.G AELSK.L CA PE=1 SV=2 AELSK.L AELSK.L AELSK.L AELSK.L AELSK.L	SV=3 2 SV=2
Prote KAPC, CAMP- KAPC, CAMP- KAPC, CAMP- KAPC,	Top scoring pept: Score greater the Score greater the Score greater the Core Expect 80.4 8.5e-06 76.9 1.9e-05 38.7 0.13 18.0 15 12.6 51 12.6 51 12.6 51 12.6 51 11.9 61	an 30 indicat: Delta Hi -0.2750 -0.2750 -0.2750 -0.2111 -0.2111 -0.2111 -0.2111	es identity t Protein 1 KAPCA_BOVIN 1 KAPCA_BOVIN 1 KAPCA_BOVIN 1 KAPCA_BOVIN 3 GSA_XYLFT 2 GSA_XYLFT 2 GSA_XYLFT 3 GSA_XYLFT	Peptide R. UTLCGTPEYLAPEIII R. TUTLCGTPEYLAPEII R. TUTLCGTPEYLAPEII R. TUTLCGTPEYLAPEII K. GGSGMLTLGTPSSPOVI K. GGSGMLTLGTPSSPOVI K. GGSGMLTLGTPSSPOVI	PRKACA PE=2 SK.G SK.G SK.G AELSK.L AELSK.L AELSK.L AELSK.L AELSK.L AELSK.L AELSK.L	SV=3 2 SV=2
Proto KAPCJ CANP- KAPCJ CANP- KAPCJ CANP- KAPCJ CANP- KAPCJ CANP-	Top scoring pept: Score greater the Score greater the Score Bayest 80.4 8.5e-06 76.9 1.9e-05 38.7 0.13 18.0 15 12.6 51 12.6 51 12.6 51 12.6 51 11.9 61	an 30 indicat. Delta Hii -0.2750 -0.2750 -0.2750 -0.2111 -0.2111 -0.2111 -0.2111 -0.2111	== homology es identity t Protein 1 KAPCA_BOVIN 1 KAPCA_BOVIN 1 KAPCA_BOVIN 3 GSA_XYLFT 3 GSA_XYLFT 2 GSA_XYLFM 3 GSA_XYLFM 3 GSA_XYLFT	Peptide R.IWILGTPEYLAPEIII R.TWILGTPEYLAPEIII R.TWILGTPEYLAPEIII R.TWILGTPEYLAPEII K.GGSCHILGIPSSPOVI K.GGSCHILGIPSSPOVI K.GGSCHILGIPSSPOVI K.GGSCHILIGIPSSPOVI	PRKACA PE=2 SK.G N=PRKACA PE=2 SK.G N=PRKACA PE=2 SK.G N=PRKACA PE=2 N=PRKACA PE=1 N=PRKACA PE=1 N=N N=N N=N N=N N=N N=N N=N N=N N=N N=	SV=3 2 SV=2
Prote KAPC. cAMP- KAPC. cAMP- KAPC. cAMP- KAPC. cAMP- KAPC. cAMP- KAPC. cAMP- KAPC.	Top scoring pept: Score greater th Score greater th Score greater th B0.4 6.5e-06 76.9 1.9e-05 38.7 0.13 18.0 15 12.6 51 12.6 51 12.6 51 12.6 51 12.6 51 12.6 51 12.9 61 11.9 61 11.9 61	an 30 indicat. an 42 indicat. Delta Hii -0.2750 -0.2750 -0.2750 -0.2750 -0.2111 -0.2111 -0.2111 -0.2111 -0.2111 -0.2111	== homology es identity t Protein 1 KAPCA_BOVIN 1 KAPCA_BOVIN 1 KAPCA_BOVIN 3 GSA_XVLFT 3 GSA_XVLFT 3 GSA_XVLFT 3 GSA_XVLFT 3 GSA_XVLFT 3 GSA_XVLFT	Peptide R.TWILGGTPEYLAPEIII R.TWILGGTPEYLAPEII R.TWILGTPEYLAPEII K.GGSGMLIGJPSSFOW K.GGSGMLIGJPSSFOW K.GGSGMLIGJPSSFOW K.GGSGMLIGJPSSFOW K.GGSGMLIGJPSSFOW K.GGSGMLIGJPSSFOW	PRKACA PE=2 SK.G SK.G N=PRKACA PE=3 SK.G AELSK.L CA PE=1 SV=3 AELSK.L CA PE=1 SV=3 AELSK.L PE=1 SV=4 (1)	SV=3 2 SV=2
Prote KAPC, CAMP- KAPC, CAMP- KAPC, CAMP- KAPC, CAMP- KAPC,	Top scoring pept: Score greater the Score greater the Score Bayeot 80.4 8.5e-06 76.9 1.9e-05 12.6 51 12.6 51 12.6 51 12.6 51 11.9 61 11.9 61 11.9 61	an 30 indicat. an 42 indicat. Delta Hi -0.2750 -0.2750 -0.2750 -0.2750 -0.2750 -0.2111 -0.211 -0.211	== homology == identity t Protein 1 KAPCA_BOVIN 1 KAPCA_BOVIN 1 KAPCA_BOVIN 1 KAPCA_BOVIN 3 GSA_XVLFT 3 GSA_XVLFT 3 GSA_XVLFT 79 Matches: 79 Matches:	Peptide R. TUTLCGTPEYLAPEIII R. TUTLCGTPEYLAPEIII R. TUTLCGTPEYLAPEIII K. GOSCMLTLCTPSSPOW K. GOSCMLTLCTPSSPOW K. GOSCMLTLCHSSPOW K. GOSCMLTLCTPSSPOW K. GOSCMLTLCTPSSPOW III Sequences II III Sequences II	PRKACA PE=2 SK.G SK.G N=PRKACA PE=2 SK.G N=PRKACA PE=2 N=PRKACA PE=1 N=PRKACA PE=1 N=PRK	5V=3 2 5V=2
Prote KAPC, CAMP- KAPC, CAMP- KAPC, CAMP- KAPC, CAMP- KAPC,	Top scoring pept: Score greater the Score greater the Score greater the Association of the Score States of	an 30 indicat. an 42 indicat. Delta Hi -0.2750 -0.2750 -0.2750 -0.2750 -0.2111 -0.211	monology sidentity t Protein 1 KAPCA_BOVIN 1 KAPCA_BOVIN 1 KAPCA_BOVIN 3 GSA_XVLFT 3 GSA_XVLFT 2 GSA_XVLFT 2 GSA_XVLFT 3 GSA_XVLFT 3 GSA_XVLFT 3 GSA_XVLFT 3 GSA_XVLFT 3 GSA_XVLFT	Peptide R.TuTLGGTPEYLAPEIII R.TuTLGGTPEYLAPEII R.TuTLGGTPEYLAPEII K.GSGMLLGPESPOW K.GSGMLLGPESPOW K.GSGMLLGPESPOW K.GSGMLLGPESPOW K.GSGMLLGPESPEW I.GSGMLLGPESPOW	PRKACA PE=2 SK.G SK.G N=PRKACA PE=3 SK.G N=PRKACA PE=1 SK.G N=PRKACA PE=1 SK.G N=PRKACA PE=2 N=PRKACA PE=3 N=PRKACA	SV=3 2 SV=2

However, casein peptides are unusually easy to analyse. Here is a more typical example of what you can expect to find - a strong match to a phosphopeptide from a protein kinase.

There is little to choose in terms of score between having the phosphate on T1 or T3.



We can see why there is little difference in score between placing the phosphate on T1 or T3. There is just one extra matched peak, and in probability terms, there isn't a huge difference between 20 matches using 55 experimental peaks and 21. However, if you had to choose one or the other, you'd probably go for T1



	or 62 ppm		Mass	s (Da)		 2
(udd)	1	÷.,				
2 -70	1					
^ነ -80	+		•			
-90	+++++++++++++++++++++++++++++++++++++++	E00	4000			
RMS err	or 62 ppm	500	1000 1500 Mas:	s (Da)		
NCBI	BLAST sea	rch of T	WILCGTPEYLAPEIILSK			
(Paran	neters: blastp	, nr prote	in database, expect=20000, no filter,	PAM30)		
Other	BLAST web	gateway	<u>s</u>			
Allma	tches to thi	s ouerv				
		1				
Score	Mr(calc)	Delta	Sequence	Site Analysis		
80.4	2214.0683	-0.2750	TWTLCGTPEYLAPEIILSK	Phospho T1 69.17%		
76.9	2214.0683	-0.2750	TWTLCGTPEYLAPEIILSK	Phospho T3 30.83%		
38.7	2214.0683	-0.2750	TWTLCGTPEYLAPEIILSK	Phospho T7 0.00%		
18.0	2214.0683	-0.2750	TWTLCGTPEYLAPEIILSK	Phospho Y10 0.00%		
12.6	2214.0044	-0.2111	GGSGMLTLGIPSSPGVPAELSK			
12.6	2214.0044	-0.2111	GGSGMLTLGIPSSPGVPAELSK			
12.6	2214.0044	-0.2111	GGSGMLTLGLPSSPGVPAELSK			
	2214.0044	-0.2111	GGSGMLTLGLPSSPGVPAELSK			
12.6	2214.0044	-0.2111	GGSGMLTLGIPSSPGVPAELSK			
12.6 11.9	2214.0044	-0.2111	GGSGMLTLGIPSSPGVPAELSK			
12.6 11.9 11.9						
12.6 11.9 11.9			Maccat: b	ttp://www.matrixscience	<u>.com/</u>	
12.6 11.9 11.9			Mascot.	-		

The delta score site analysis suggests 70% probability on T1 and 30% on T3 ... much less clear cut. We can't be confident which site is modified, or whether there is a mixture of both isoforms. But, we can be confident it is not on T7 or Y10 because the score drops dramatically, and these are assigned 0% probability.

Sometimes, it is worth looking at the sequence annotations to see whether these are known phosphorylation sites. If the database sequence doesn't have detailed annotations, you can follow the BLAST link to try and match the peptide to an entry from a better annotated database. In this case, we're searching SwissProt, so we can go straight to the protein view report

<pre>Fi NOU_RED 19 19 19 100 photocollate (by Shiftar Ky). Fi NOU_RES 100 140 Phosphostine (by Shiftar Ky). FI NOU_RES 100 140 Phosphostine (by Shiftar Ky). FI NOU_RES 100 140 Phosphostine. FI NOU_RES 100 19 Phosphostine. FI NOU_RES 200 200 Phosphostine. FI CONFLICT 200 200 F -> N (in Ref. 4; AA sequence). FI CONFLICT 200 200 F -> N (in Ref. 4; AA sequence). FI CONFLICT 200 200 F -> N (in Ref. 4; AA sequence). FI CONFLICT 200 200 F -> N (in Ref. 4; AA sequence). FI CONFLICT 200 200 F -> N (in Ref. 4; AA sequence). FI CONFLICT 200 200 F -> N (in Ref. 4; AA sequence). FI CONFLICT 200 200 F -> N (in Ref. 4; AA sequence). FI CONFLICT 200 200 F -> N (in Ref. 4; AA sequence). FI CONFLICT 200 200 F -> N (in Ref. 4; AA sequence). FI CONFLICT 200 200 F -> N (in Ref. 4; AA sequence). FI CONFLICT 200 200 F -> N (in Ref. 4; AA sequence). FI CONFLICT 200 200 F -> N (in Ref. 4; AA sequence). FI T STRAND 54 63 FI THELIX 16 32 FI THELIX 16 32 FI THELIX 16 14 143 FI STRAND 144 52 FI THELIX 16 14 143 FI STRAND 144 52 FI THELIX 129 136 FI THELIX 129 136 FI THELIX 129 136 FI THELIX 200 201 FI THELIX 200 201 FI THELIX 200 201 FI THELIX 200 201 FI THELIX 200 203 FI THELIX 200 203 FI THELIX 200 204 FI THELIX 200 204 FI THELIX 200 205 FI THELIX 200</pre>	<pre>Find Decks 140 From Find Phosphoet Control (by Sharks typ). Find Decks 140 From Find Phosphoet (by Sharks typ). Find Decks 150 150 Find Phosphoet (by Sharks typ). Find Decks 120 20 From Find Phosphoet (by Sharks typ). Find Decks 120 20 20 From Find Phosphoet (by Sharks typ). Find Decks 120 20 20 20 From Find Phosphoet (b) find Find Find Find Find Find Find Find F</pre>	4	⇒ C 🖷	🛇 www.	matrixscien	ce.com/cgi/protein_view.pl?file=%2Fdata%2F20120704%2FFtGmIfewT.dat&hit=KAPCA_BOVIN&db_ 🏠	3
if NOT_RES 196 Phosphothreonine (by similarity). = T1 FT NOT_RES 196 Phosphothreonine (by similarity). = T3 FT NOT_RES 202 202 Phosphothreonine (by similarity). = T3 FT NOT_RES 303 Phosphothreonine (by similarity). = T3 FT NUTAEN 3 N-bytistoylation. FT NUTAEN 3 N-bytistoylation. FT COMFLICT 202 202 T -> N (in Ref. 4; Ak sequence). FT COMFLICT 204 E -> Q (in Ref. 4; Ak sequence). FT COMFLICT 205 205 L -> S (in Ref. 4; Ak sequence). FT TC COMFLICT 207 N -> D (in Ref. 4; Ak sequence). FT TELIX 14 43 FT TELIX 14 43 FT STEAMD 54 63 FT STEAMD 14 12 FT STEAMD 141 12 FT STEAMD 141 12	<pre>if woo-EES 196 196</pre>	FT	MOD DES	140	140	Phosphochical (by Similarity).	2
<pre>Fit NOD_PES is0 100 Phosphothreonime.(by PDFK1. FT NOD_PES 302 202 Phosphothreonime.(by PDFK1. FT NOD_PES 303 339 Phosphoterine. FT NUTACEN 3 3 N->D: No myristoylation. FT NUTACEN 3 3 N->D: No myristoylation. FT CONFLICT 202 202 T -> N (in Ref. 4; Ak sequence). FT CONFLICT 204 204 E -> Q (in Ref. 4; Ak sequence). FT CONFLICT 204 206 L -> S (in Ref. 4; Ak sequence). FT CONFLICT 206 206 L -> S (in Ref. 4; Ak sequence). FT CONFLICT 207 287 N -> D (in Ref. 2; Ak sequence). FT CONFLICT 208 287 N -> D (in Ref. 2; Ak sequence). FT STRAND 44 52 FT STRAND 54 63 FT STRAND 54 63 FT STRAND 54 63 FT STRAND 54 63 FT STRAND 107 112 FT STRAND 114 122 FT STRAND 114 122 FT STRAND 114 123 FT STRAND 114 124 FT STRAND 114 125 FT STRAND 114 125 FT STRAND 114 126 FT HELIX 208 203 FT HELIX 208 203</pre>	<pre>FT HOUTRES 198 198 FPGsphothreoning: by PDFH1. ')' = T3 FT HOUTRES 202 202 Phosphothreoning (By similarity). = T3 FT HOUTRES 339 339 Phosphothreoning (By similarity). = T3 FT HOUTREN 3 3 N->D: No myristoylation. FT CONFLICT 204 204 E -> 0 (in Ref. 4; AA sequence). FT CONFLICT 205 205 T -> N (in Ref. 4; AA sequence). TT CONFLICT 205 206 L -> 0 (in Ref. 4; AA sequence). TT CONFLICT 206 206 L -> 0 (in Ref. 4; AA sequence). TT CONFLICT 207 207 N -> D (in Ref. 4; AA sequence). TT CONFLICT 208 206 N -> D (in Ref. 4; AA sequence). TT CONFLICT 208 206 L -> 0 (in Ref. 4; AA sequence). TT TUNN 644 66 FT STRAND 44 52 FT HELIX 11 43 FT HELIX 11 43 FT HELIX 11 43 FT HELIX 129 203 TT STRAND 54 63 FT HELIX 129 123 FT HELIX 203 205 FT HELIX 129 123 FT HELIX 203 205 FT HELIX 203 205 FT HELIX 204 204 FT HELIX 203 205 FT HELIX 203 205 FT HELIX 204 204 FT HELIX 203 205 FT HELIX 203 307 FT HELIX 204 204 FT HELIX 204 204 FT HELIX 204 204 FT HELIX 204 204 FT HELIX 204 205 FT HELIX 206 211 FT HELIX 206 214 FT HELIX 206 214 FT HELIX 206 214 FT HELIX 206 214 FT HELIX 206 203 FT HELIX 206 204 FT HELIX 206 204 FT HELIX 206 205 FT HELIX 206 FT HELIX 206 205 FT HELIX 206 FT HELIX 206</pre>	FT	MOD_RES	196	196	Phosphothraphics (By similarity) - T1	_
<pre>Find box bits is a box box provide (by similarity). = T3 Find box bits is a box box provide (by similarity). = T3 Find box box bits is a box box bits box bits box box box box box box box box box box</pre>	<pre>11 BIO 7E2 352 152 150 FROM DUPLICATION DY FURT. 11 BIO 7E2 353 339 FROM DUPLICATION DY FURT. 11 BIO 7E2 339 FROM DUPLICATION DY FURT. 11 BIO 7E2 339 FROM DUPLICATION DY FURT. 11 BIO 7E2 353 SUBJECT STATES 11 CONFLICT 202 202 T -> N (in Ref. 4; Ak sequence). 11 CONFLICT 204 204 L -> Q (in Ref. 4; Ak sequence). 11 CONFLICT 206 206 L -> S (in Ref. 4; Ak sequence). 11 CONFLICT 207 N -> D (in Ref. 4; Ak sequence). 11 CONFLICT 207 Q 207 N -> D (in Ref. 4; Ak sequence). 11 CONFLICT 207 Q 207 N -> D (in Ref. 4; Ak sequence). 11 CONFLICT 207 Q 207 N -> D (in Ref. 4; Ak sequence). 11 CONFLICT 207 Q 207 N -> D (in Ref. 4; Ak sequence). 11 CONFLICT 207 Q 207 N -> D (in Ref. 4; Ak sequence). 11 STRAND 44 53 11 STRAND 44 63 11 STRAND 54 66 11 STRAND 107 112 11 STRAND 107 112 11 STRAND 107 112 11 STRAND 107 113 11 STRAND 114 123 11 HELIX 209 233 11 HELIX 209 234 11 HELIX 209 234 11 HELIX 209 234 11 HELIX 209 235 11 HELIX 209 235 12 HELIX 209 235 13 HELIX 209 235 14 HELIX 209 235 15 HELIX 209 240 HEL</pre>	F I FT	MOD_RES	100	100	Phosphotheonine (by Shillarity).	
/ HOU RLS 202 202 Phosphoterionine (by Similarity) + + + + + + + + + + + + + + + + + +	<pre>// HOUPAGE 202 202 Prosphormeon.he (by Similarity)</pre>	F 1	NOD_RES	198	198	Phosphothreenine; by $PDrk1$. = T3	
<pre>Fi NOU ALS 339 339 FindBindSetInE. Fi LUFID 2 N=Wristoy1glyCinE. FT COMFLICT 202 202 N=>>N (mpRi) (ylpCinE. FT COMFLICT 204 204 E => (in Ref. 4) AA sequence). FT COMFLICT 204 206 L => S (in Ref. 4) AA sequence). FT COMFLICT 207 207 N => D (in Ref. 4) AA sequence). FT COMFLICT 207 207 N => D (in Ref. 4) AA sequence). FT COMFLICT 207 208 208 L => S (in Ref. 4) AA sequence). FT COMFLICT 208 208 L => S (in Ref. 4) AA sequence). FT COMFLICT 208 208 L => S (in Ref. 4) AA sequence). FT COMFLICT 208 208 L => S (in Ref. 4) AA sequence). FT STRAND 54 63 FT STRAND 54 63 FT STRAND 54 63 FT STRAND 107 112 FT HELIX 108 696 FT STRAND 107 112 FT HELIX 129 136 FT STRAND 107 112 FT HELIX 129 136 FT HELIX 141 160 FT HELIX 129 136 FT HELIX 208 211 FT HELIX 208 211 FT HELIX 208 211 FT HELIX 208 201 FT HELIX 208 203 FT HELIX 303 307 FT HELIX 303 307 FT HELIX 304 307 FT HELIX 305 307 FT HELIX 305 307 FT HELIX 306 307 FT HELIX 209 233 FT HELIX 140 CON FY SPDD227D2DEFESD CRC54! SGMALAKKG SRCSWKFL AKAREPILKK WENPLONTAH LOVERINTL GYSGROWL SGMALAKKG SRCSWKFL AKAREPILKK WENPLONTAH LOVERINTL GYSGROWL SINGKKFL AKAREPILKK WENPLONTAH SUFFRENCES F</pre>	<pre>FI NOU RES 333 333 730 Prosposerine: FI LUTED 3 1 N-D: No mytacoyl giycine. FT CONFLICT 204 204 F -> 0 (in Ref. 4; Ak sequence). FT CONFLICT 206 206 L -> 8 (in Ref. 4; Ak sequence). FT CONFLICT 206 206 L -> 8 (in Ref. 4; Ak sequence). FT CONFLICT 207 N -> D (in Ref. 4; Ak sequence). FT CONFLICT 207 N -> D (in Ref. 4; Ak sequence). FT T TURN 64 66 FT STRAND 44 52 FT HELIX 16 726 FT STRAND 54 63 FT STRAND 144 122 FT HELIX 167 76 FT STRAND 144 122 FT HELIX 167 76 FT STRAND 144 122 FT HELIX 167 76 FT STRAND 173 175 FT STRAND 173 175 FT STRAND 181 183 FT HELIX 203 205 FT HELIX 204 273 FT HELIX 204 273 FT HELIX 205 211 FT HELIX 205 211 FT HELIX 206 211 FT HELIX 206 211 FT HELIX 206 211 FT HELIX 206 211 FT HELIX 207 175 FT STRAND 181 183 FT HELIX 208 211 FT HELIX 208 211 FT HELIX 209 209 FT HELIX</pre>	P I	MOD_RES	202	202	Phosphothreonine (By Similarity).	
<pre>FI LP1D 2 2 2 N=myristoyl glycine. FT LVTACEN 3 N→D: No myristoyleton. TT CONFLICT 20 202 T→N (in Ref. 4; AA sequence). FT CONFLICT 20 20 0 L→S (in Ref. 4; AA sequence). FT CONFLICT 20 20 0 L→S (in Ref. 4; AA sequence). FT CONFLICT 20 20 0 L→S (in Ref. 4; AA sequence). FT CONFLICT 20 20 0 L→S (in Ref. 4; AA sequence). FT CONFLICT 20 20 0 L→S (in Ref. 4; AA sequence). FT CONFLICT 20 20 0 L→S (in Ref. 2; AA sequence and 3; AA sequence). FT HELIX 41 43 FT STRAND 54 63 FT TURN 64 66 FT HELIX 77 52 FT STRAND 69 76 FT HELIX 86 96 FT HELIX 10 10 112 FT STRAND 10 112 FT HELIX 20 205 FT HELIX 203 205 FT HELIX 203 205 FT HELIX 203 205 FT HELIX 204 233 FT HELIX 204 253 FT HELIX 206 231 FT HELIX 206 235 FT HELIX 206 235</pre>	<pre>Ff L1P1D 2 2 2 N-myristoyl glychne. FT NUTACEN 3 3 N-myristoyl glychne. TT COMFLICT 202 202 T → N (in Ref. 4; AA sequence). TT COMFLICT 204 204 L → O (in Ref. 4; AA sequence). FT COMFLICT 205 204 L → O (in Ref. 4; AA sequence). FT COMFLICT 205 204 L → O (in Ref. 4; AA sequence). FT COMFLICT 207 207 S = S = S = S = S = S = S = S = S = S</pre>	FT	MOD_RES	339	339	Phosphoserine.	
FT NUTACEM 3 3 N-D: No myristoylation. FT CORFLICT 202 Z - N (in Ref. 4; Ak sequence). FT CORFLICT 203 204 E \rightarrow S (in Ref. 4; Ak sequence). FT CORFLICT 208 207 N \rightarrow N (in Ref. 4; Ak sequence). FT CORFLICT 208 207 N \rightarrow D (in Ref. 2; Ak sequence and 3; Ak FT CORFLICT 208 207 N \rightarrow D (in Ref. 2; Ak sequence and 3; Ak FT HELIX 14 32 FT HELIX 14 43 FT HELIX 14 44 FT STRAND 54 63 FT STRAND 54 66 FT STRAND 64 76 FT STRAND 107 112 FT HELIX 196 96 FT STRAND 107 112 FT HELIX 129 136 FT STRAND 114 122 FT HELIX 141 160 FT HELIX 141 160 FT HELIX 142 153 FT HELIX 209 205 FT STRAND 173 175 FT STRAND 181 183 FT HELIX 209 205 FT HELIX 209 205 FT HELIX 209 205 FT HELIX 200 20	<pre>FT NUTACEN 3 3 N→D: No mytistoylation. TT COMFLICT 202 202 T → N (in Ref. 4; AA sequence). TT COMFLICT 204 204 E → O (in Ref. 4; AA sequence). TT COMFLICT 207 207 N→D (in Ref. 4; AA sequence). TT COMFLICT 207 207 N→D (in Ref. 4; AA sequence). TT COMFLICT 207 207 N→D (in Ref. 4; AA sequence). TT COMFLICT 207 207 N→D (in Ref. 4; AA sequence). TT COMFLICT 207 207 N→D (in Ref. 4; AA sequence). TT COMFLICT 207 207 N→D (in Ref. 4; AA sequence). TT COMFLICT 207 207 N→D (in Ref. 4; AA sequence). TT TURN 141 43 TT TURN 64 66 TT STRAND 54 63 TT STRAND 54 63 TT STRAND 54 63 TT HELIX 77 66 TT HELIX 170 172 TT HELIX 141 160 TT HELIX 141 160 TT HELIX 141 160 TT HELIX 203 205 TT HELIX 203 307 TT HELIX 203 307 TT HELIX 204 273 TT TURN 266 289 FT HELIX 296 298 FT HELIX 296 298 FT HELIX 296 298 FT HELIX 303 307 FT HELIX 346 348 S0 SEQUENCE 351 AA: 40620 HW; S9DDD227D2DEEESD CRC64; NGHAAKKG SPCSWERL AKAEPJEKK WENPANTAH LDC/FERIKIL GTOSFGRWHL NGHAKKKG SPCSWERL AKAEPJEKK WENPANTAH LDC/FERIKIL GTOSFGRWHL NGHAKKG SPCSWERL AKAEPJEKK WENPANTAH LDC/FERIKIL GTOSFGRWHL NGHAKHKG SPCSWERL AKAEPJEKK WENPANTAH LDC/FERIKILON TY NGHAKHKG SPCSWERL AKAEPJEKK WENPANTAH LDC/FERIKIL</pre>	FT	LIPID	2	2	N-myristoyl glycine.	
$\begin{array}{llllllllllllllllllllllllllllllllllll$	<pre>FT COMFLICT 202 202 T → N (in Ref. 4; AA sequence). FT COMFLICT 204 204 L → O (in Ref. 4; AA sequence). FT COMFLICT 205 206 L → S (in Ref. 4; AA sequence). FT COMFLICT 207 N → D (in Ref. 4; AA sequence). FT COMFLICT 207 N → D (in Ref. 4; AA sequence). FT COMFLICT 207 N → D (in Ref. 4; AA sequence). FT STEAND 54 63 FT STEAND 54 63 FT STEAND 54 63 FT STEAND 54 63 FT STEAND 107 112 FT STEAND 107 112 FT HELIX 107 82 FT STEAND 114 122 FT HELIX 107 125 FT STEAND 114 122 FT HELIX 203 205 FT HELIX 203 205 FT HELIX 203 205 FT HELIX 204 273 FT HELIX 204 273 FT HELIX 204 253 FT HELIX 205 298 FT HELIX 206 298 FT HELIX 303 307 FT HELIX 141 COMPACES CRC64; NGMAAKKG SUGENCEL AAREDFLEK WENPAONTAH LOOFERINTL GESFERVEL NGMAAKKG SUGENCEL AAREDFLEK WENPAONTAH LOOFERINTL FERSTROWL NGMAAKKG SUGENCEL AAREDFLEK WENPAONTAH LOOFERINTL FERSTROWL NGMAAKKG SUGENCEL AFAREDFLEK WENPAONTAH LOOFERINTLY WENPAONTAH LOOFER NGMAAKKG SUGENCEL AFAREDFLEK WENPAONTAH LOOFERINTLY FERSTROWEN NGMAAKKG SUGENCEL AFAREDFLEK WENPAONTAH LOOFERINTLY FERSTROWEN NGMAAKKG SUGENCEL AFAREDFLEK WENPAONTAH LOOFERINTLY FERSTROWEN NGMAAKKG SUGENCEL AFAREDFLEK WENPAONTAH LOOFERINT FERSTROWENPFF ADOPTOFFF ANNON AFAREDFLEKK WENPAONT</pre>	FT	MUTAGEN	3	3	N->D: No myristoylation.	
$\begin{array}{llllllllllllllllllllllllllllllllllll$	<pre>FT COMFLICT 204 204 E → 0 (in Ref. 4; AA sequence). FT COMFLICT 205 206 L → S (in Ref. 4; AA sequence). FT COMFLICT 205 207 N → D (in Ref. 2; AA sequence and 3; AA sequence). FT ELLIX 41 43 FT HELIX 41 45 FT STRAND 54 63 FT STRAND 54 66 FT STRAND 69 76 FT HELIX 77 82 FT HELIX 86 96 FT STRAND 107 112 FT STRAND 107 112 FT STRAND 107 172 FT HELIX 101 102 FT HELIX 205 205 FT HELIX 208 211 FT HELIX 208 211 FT HELIX 208 214 FT HELIX 208 215 FT HELIX 208 214 FT HELIX 208 21</pre>	FT	CONFLICT	202	202	T -> N (in Ref. 4; AA sequence).	
<pre>FT CONFLICT 206 206 L -> S (in Ref. 4; AA sequence). FT CONFLICT 207 207 N -> D (in Ref. 2; AA sequence). FT FT CONFLICT 207 N -> D (in Ref. 2; AA sequence and 3; AA sequence). FT STEAMD 44 43 FT STEAMD 54 453 FT TUEN 54 55 FT STEAMD 69 76 FT HELIX 77 62 FT STEAMD 107 112 FT HELIX 186 96 FT STEAMD 107 112 FT HELIX 129 136 FT STEAMD 107 112 FT HELIX 141 160 FT HELIX 141 160 FT HELIX 141 160 FT HELIX 203 205 FT STEAMD 83 183 FT HELIX 203 205 FT HELIX 203 205 FT HELIX 204 253 FT HELIX 206 201 FT HELIX 206 201 FT HELIX 206 201 FT HELIX 206 201 FT HELIX 206 203 FT HELIX 141 100 FT HELIX 206 203 FT HELIX 206 203 FT HELIX 206 203 FT HELIX 141 100 FT HELIX 141 100 FT HELIX 206 203 FT HELIX 206 203 FT HELIX 206 203 FT HELIX 141 100 FT HELIX 206 203 FT HELIX 206 203 FT HELIX 206 203 FT HELIX 141 100 FT HELIX 206 203 FT HELIX 206 203 FT HELIX 206 203 FT HELIX 206 203 FT HELIX 141 100 FT HELIX 206 203 FT HELIX 206 FT HELIX 206 FT HELIX FT HEAPATAH LOPERITL FT HEAPATAH H</pre>	<pre>FT CONFLICT 206 206 L → S (in Ref. 4; AA sequence). FT CONFLICT 287 N → D (in Ref. 4; AA sequence). FT CONFLICT 287 287 N → D (in Ref. 4; AA sequence and 3; AA sequence). FT HELIX 16 32 FT HELIX 14 43 FT STRAND 44 63 FT STRAND 54 66 FT STRAND 54 66 FT STRAND 107 112 FT HELIX 86 96 FT STRAND 114 122 FT HELIX 141 160 FT HELIX 295 234 FT HELIX 296 298 FT HELIX 303 307 FT HELIX 346 348 SO SQUENCE 351 AA; SODDAZ27DZDEEESD CRC64; NGMAAKKG SUCSYNCH AKARDPIKK WENPANTAR LDOFERIKTL GTOSTORVEL NGMAAKKG SUCSYNCH AKARDPIKK WENPANTAR LDOFERIKTL FTOSTORVEL NGMAAKKG SUCSYNCH HEDDAKKEL NHAD VUNCUUT EMACYPYF ADOPTOTYER NVGGYFFF HEDDAKKEL NHAD VUNCUUT EMACYPYFF ADOPTOTYER NVGGYFFF HEDDAKKEL NHAD VUNCUUT EMACYPYFF ADOPTOTYER NVGGYFFF HEDDAKKEL NHAD VEEELEVSI NEKCGKEFSE F</pre>	FT	CONFLICT	204	204	E -> Q (in Ref. 4; AA sequence).	
$ \begin{array}{llllllllllllllllllllllllllllllllllll$	FT CONFLICT 287 287 N → D (in Ref. 2; AA sequence and 3; AA sequence). FT HELIX 16 32 FT HELIX 41 43 FT STRAND 44 52 FT STRAND 54 63 FT STRAND 64 66 FT STRAND 144 52 FT STRAND 144 52 FT STRAND 144 52 FT STRAND 114 122 FT STRAND 114 123 FT HELIX 170 172 FT STRAND 181 183 FT HELIX 203 205 FT HELIX 203 205 FT HELIX 204 273 FT STRAND 181 183 FT HELIX 244 253 FT HELIX 296 299 FT HELIX 296 299 FT HELIX 396 390 FT HELIX 396 390 FT HELIX 396 299 FT HELIX 396 299 FT HELIX 1466 0 NF; SPDDD227D2DEEESD CRC64; KMALALKEG SEGSVKFL AKAELFLKK WENPAONTAM LOGFERIKIL GTOSFGRVHL VEHETOMY MELLORGV VELOPIELIK NEWPAONTAM LOGFERIKIL GTOSFGRVHL VEHETOMY MELLORGV VELOPIELIK WENPAONTAM LOGFERIKIL GTOSFGRVHL VEHETOMY ANGLORGV VELOPIELI HERFANDAW PFILVERFE FKNNSNLYMV MEVVPOCFFF SHLRENGER EMALEVIAL UNIT FYLHS LDLITEDLEF FKNLIDQOOY LOUTDFORK RVKGTHER INSOUNDED VELEELEVISI NECCOKEFSE F	FT	CONFLICT	206	206	L -> S (in Ref. 4; AA sequence).	
Sequence). FT HELIX 16 32 FT HELIX 41 43 FT HELIX 41 43 FT STEAND 44 52 FT STEAND 54 63 FT STEAND 54 63 FT TTOM 64 63 FT TTOM 64 63 FT TTOM 64 63 FT TTOM 64 63 FT TELIX 77 82 FT HELIX 107 112 FT STEAND 107 112 FT HELIX 141 160 FT HELIX 141 160 FT STEAND 173 175 FT STEAND 131 181 FT HELIX 203 211 FT HELIX 204 233 FT	FT December 1997 FT HELIX 16 32 FT HELIX 41 43 FT HELIX 41 43 FT STRAND 54 63 FT STRAND 54 63 FT STRAND 54 63 FT STRAND 19 66 FT STRAND 19 76 FT STRAND 107 112 FT HELIX 166 96 FT STRAND 107 112 FT HELIX 129 136 FT STRAND 107 112 FT HELIX 129 136 FT STRAND 173 175 FT STRAND 181 183 FT HELIX 200 211 FT HELIX 200 211 FT HELIX 200 211 FT HELIX 200 211 FT HELIX 244 253 FT HELIX 290 293 FT HELIX 296 298 FT HELIX 296 298 FT HELIX 303 307 FT HELIX 346 348 S0 SCUENCE 351 A&: 40620 HW; 59DDD227D2DEEE5D CRC64; KGNAALKKS SECSEVERL & KAREDFLEK WENPAQNTAH LD0FERIKTL GTOSFGNWEL KGNAALKKS SECSEVERL & KAREDFLEK WENPAQNTAH LD0FERIKTL FTOSFGNWEL KGNAALKKS SECSEVERL & KAREDFLEK WENPAQNTAH LD0FERIKTL FENALWYN VENHETONNY WENFEN WENPFFF AUDPTOJVER WENFFF HESOFKEL KAREDFLEK WENPAQNTAH LD0FERIKTL FENALWYN VENHETONNY WENFFFF HESOFKEL KAREDFLEK WENPAQNTAH LD0FFF FINANNING VENHETONNY WENFFF HESOFKEL WENPAQNTAH LD0FFF FINANNING VENHETONNY WENFFF HESOFKEL WENPAQNTAH LD0FFF FINANNING VENHETONNY WENFFF HESOFKEL WENFFF FINAN	FT	CONFLICT	287	287	N -> D (in Ref. 2; AA sequence and 3; AA	
FT HELIX 16 32 FT HELIX 41 43 FT STRAND 44 52 FT STRAND 54 63 FT STRAND 64 66 FT STRAND 69 76 FT TURN 64 66 FT HELIX 77 82 FT HELIX 77 82 FT HELIX 100 112 FT FELIX 141 160 FT HELIX 170 172 FT HELIX 170 172 FT STRAND 173 175 FT STRAND 173 175 FT HELIX 203 205 FT HELIX 203 205 FT HELIX 204 233 FT HELIX 204 234 FT HELIX 206 233 FT HELIX 206 235 FT	FT HELIX 16 32 FT HELIX 41 43 FT STRAND 44 52 FT STRAND 54 63 FT TURN 64 66 FT HELIX 77 82 FT STRAND 107 112 FT HELIX 86 96 FT HELIX 177 82 FT STRAND 107 112 FT STRAND 107 112 FT STRAND 107 112 FT STRAND 107 112 FT STRAND 113 126 FT HELIX 208 205 FT HELIX 203 205 FT HELIX 303 307 FT HELIX 304 303 FT HELIX 304 4253 FT HELIX 304 456 FT HELIX 506 VN; SPDD227D2DEEE5D CRC64/ FT HELIX 304 306 FT HELIX 100 90 FT HEL	FΤ				sequence).	
FT HELIX 41 43 FT STRAND 44 52 FT STRAND 54 63 FT STRAND 64 66 FT STRAND 69 76 FT STRAND 69 76 FT STRAND 69 76 FT STRAND 69 76 FT STRAND 107 12 FT STRAND 107 12 FT STRAND 107 12 FT STRAND 107 12 FT STRAND 101 12 FT STRAND 101 12 FT STRAND 13 160 FT HELIX 201 201 FT STRAND 131 183 FT STRAND 131 183 FT HELIX 203 201 FT HELIX 203 201 FT HELIX 204 233 FT <td>FT HELIX 41 43 FT STRAND 44 52 FT STRAND 54 63 FT STRAND 54 63 FT STRAND 69 76 FT STRAND 69 76 FT STRAND 69 76 FT STRAND 107 112 FT HELIX 77 626 FT HELIX 170 172 FT HELIX 129 136 FT HELIX 141 160 FT HELIX 170 172 FT HELIX 203 205 FT HELIX 204 273 FT STRAND 181 183 ST SEQUENCE 351 AJ: 40620 HW; 59DDD227D2DEEE5D CRC64; NGNAALKKS S205WKFFL AKAEDFLKK WENPAQNTAH LD0FERIKTL GTOSFGKWHL NGNAALKKS S205WKFFL AKAEDFLKK WENPAQNTAH LD0FOK NGNAALKKS S205WKFFL AKAEDFLKK WENPAQNTAH LD0FFGKKKT TUDJUALYOFK WENPFFKK GFGDTSNFDD YEEEFINNS NKWOK DINNHKWFAT TUDJUALYOFK WEAFFFFKK GFGDTSNFDD YEEEFINSI NKKCKKFFSE F</td> <td>FΤ</td> <td>HELIX</td> <td>16</td> <td>32</td> <td></td> <td></td>	FT HELIX 41 43 FT STRAND 44 52 FT STRAND 54 63 FT STRAND 54 63 FT STRAND 69 76 FT STRAND 69 76 FT STRAND 69 76 FT STRAND 107 112 FT HELIX 77 626 FT HELIX 170 172 FT HELIX 129 136 FT HELIX 141 160 FT HELIX 170 172 FT HELIX 203 205 FT HELIX 204 273 FT STRAND 181 183 ST SEQUENCE 351 AJ: 40620 HW; 59DDD227D2DEEE5D CRC64; NGNAALKKS S205WKFFL AKAEDFLKK WENPAQNTAH LD0FERIKTL GTOSFGKWHL NGNAALKKS S205WKFFL AKAEDFLKK WENPAQNTAH LD0FOK NGNAALKKS S205WKFFL AKAEDFLKK WENPAQNTAH LD0FFGKKKT TUDJUALYOFK WENPFFKK GFGDTSNFDD YEEEFINNS NKWOK DINNHKWFAT TUDJUALYOFK WEAFFFFKK GFGDTSNFDD YEEEFINSI NKKCKKFFSE F	FΤ	HELIX	16	32		
FT STRAND 44 52 FT STRAND 54 63 FT STRAND 64 66 FT STRAND 69 76 FT STRAND 69 76 FT HELIX 77 62 FT HELIX 86 96 FT STRAND 107 112 FT STRAND 173 175 FT HELIX 203 205 FT HELIX 203 205 FT HELIX 204 233 FT HELIX 204 234 FT HELIX 205 307 FT HELIX 206 307 FT HELIX 206 307 FT HELIX 206 307 FT	FT STRAND 44 52 FT STRAND 54 63 FT TURN 64 66 FT HELIX 77 82 FT STRAND 107 112 FT STRAND 114 122 FT STRAND 114 122 FT STRAND 114 122 FT HELIX 129 136 FT HELIX 203 205 FT HELIX 204 273 FT HELIX 204 273 FT HELIX 304 307 FT HELIX 304 307 FT HELIX 199 293 SUBJORCE 31 AJ, 40650 NW; SUDD27D2DEEE5D CRC64: MEWAGAKENE S1 AJ, 40650 NW; SUDD27D2FEEFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF	FΤ	HELIX	41	43		
FT STRAND 54 63 FT STRAND 69 66 FT STRAND 69 76 FT HELIX 77 62 FT HELIX 86 96 FT STRAND 107 112 FT STRAND 114 122 FT HELIX 123 166 FT HELIX 141 160 FT HELIX 170 172 FT STRAND 131 160 FT HELIX 203 205 FT HELIX 203 205 FT HELIX 203 205 FT HELIX 204 231 FT HELIX 204 233 FT HELIX 296 283 FT HELIX 346 346 Stoutower 314, 40620 MU SpotD227D2DEEESD CRC64; MGRAAAKKG SEQESVKEFL AKAKEDFLKK WENAPOHAL HULEY FLMS LDLIVPER FINDUNUK FORMENAWER MUNAAKKG SEQESVKEFL AKAKEDFLKK WE	FT STRAND 54 63 FT STRAND 69 76 FT STRAND 69 76 FT STRAND 69 76 FT STRAND 69 76 FT STRAND 107 12 FT HELIX 77 82 FT STRAND 117 112 FT STRAND 117 112 FT STRAND 117 112 FT STRAND 113 123 FT STRAND 131 183 FT HELIX 208 211 FT HELIX 208 211 FT HELIX 208 211 FT HELIX 208 211 FT HELIX 244 253 FT HELIX 244 253 FT HELIX 244 253 FT HELIX 296 293 FT HELIX 296 293 FT HELIX 296 293 FT HELIX 303 307 FT HELIX 346 348 S0 SEQUENCE 351 AJ: 40620 HW; 59DDD227D2DEEESD CRC64; MGNAAAKKS SEQUENCEL ASALEPJLKK WEMPAONTAM LD0FERIKTL GTOSFGKWML WETVF0GENF SURANGE INMERCENT MERKHLOWN FFLUKTES FKNNNLYW VKHETONHY AKKLOKOK VULKQIENTL MERKHLOWN FFLUKTES FKNNNLYW VKHETONHY AKKLOKOV VELKQIENTL MERKHLOWN FFLUKTES FKNNNLYW VKHETONHY AKKLOKOV VELKARET LANGEN FFLUKTES FKNNNLYW VKHETONHY AKKLOKOV VELKARET FLUKTES FKNNNLYW VKHETONHY AKKLOKOV VELKARET FLUKTES FKNNNLYW VKHETONHY AKKO SUM FFLUKTES FKNNNLYW VKHETONHY AKKO SUM FFLUKTES FKNNNLYW VKHETONHY AKKO FUKTES FKNNNLYW VKHETONHY AKKO FUKTES FKNNNLYW VKHETONHY AKKO FUKTES FKNNNLYW FKNNNCHTON FKNNNLYW FKNNNCHTON FKNNNLYW FKNNCHTON FK	FT	STRAND	44	52		
FT TURN 64 66 FT STEAND 69 76 FT HELIX 77 62 FT STEAND 107 112 FT STEAND 107 112 FT STEAND 107 112 FT STEAND 114 122 FT STEAND 114 122 FT HELIX 129 136 FT HELIX 101 170 FT STEAND 131 183 FT STEAND 131 183 FT HELIX 203 205 FT HELIX 203 205 FT HELIX 204 233 FT HELIX 290 290 FT HELIX 290 290 FT HELIX 290 290 STEOURCH 313.41 40520 MV SPDD227D2DEEESD CRC64/ MCMAAAKG SEQESVKFL AKAEDPLKK WENNDAND FUNCHTAND FUNCHTAND FUNCHTAND MCMAAAKG SEQESVKFL AKAEDPLKK WEND FUNCHTAN	FT TUEN 64 66 FT HELIX 77 82 FT HELIX 77 82 FT HELIX 86 94 FT STRAND 107 112 FT HELIX 106 94 FT STRAND 107 112 FT HELIX 129 136 FT STRAND 114 122 FT HELIX 129 136 FT STRAND 173 175 FT HELIX 141 160 FT HELIX 170 172 FT STRAND 133 105 FT HELIX 208 211 FT HELIX 244 253 FT HELIX 244 253 FT HELIX 244 253 FT HELIX 290 293 FT HELIX 290 293 FT HELIX 303 307 FT HELIX 346 3482 FT HELIX 346 3482 FT HELIX 199 294 FT HELIX 199 295 FT HELIX 199 2	FT	STRAND	54	63		
T STEAND 69 76 T HELIX 77 82 FT HELIX 86 96 FT STEAND 107 112 FT STEAND 114 122 FT HELIX 141 124 FT STEAND 114 122 FT HELIX 141 160 FT HELIX 170 172 FT STEAND 173 175 FT STEAND 181 183 FT HELIX 208 205 FT HELIX 208 211 FT HELIX 244 253 FT HELIX 264 273 FT HELIX 264 263 FT HELIX 296 283 FT HELIX 346 348 SEQUENCY SAUGNOV VKINETONOV FFWINENUM MGNAAAKKO SEQESVKEFL AKAKEDFLKK WENNANTAH LOPERISTL GTOSFGRVHL MGNAAAKKO SEQESVKEFL AKAKEDFLKK WENNOV FYNKUV FYNKUV FNANTAH LOPUNGV	FT STEAND 69 76 FT HELIX 77 82 FT HELIX 86 96 FT HELIX 86 96 FT HELIX 120 112 FT STEAND 107 112 FT STEAND 114 122 FT STEAND 114 122 FT HELIX 129 136 FT HELIX 141 160 FT HELIX 170 172 FT STEAND 181 183 FT HELIX 203 205 FT HELIX 203 205 FT HELIX 203 205 FT HELIX 244 253 FT HELIX 244 253 FT HELIX 244 253 FT HELIX 396 299 FT HELIX 40620 NH; S9DDD227D2DEEE5D CPC64; NGRAAAKKG SECESVKEFL AKAEDFLEK WENPAONTAH LOOFERIKTL GTOSFGRVHL NGRAAAKKG SECESVKEFL AKAEDFLEK WENPAONTAH LOOFERIKTU FUNACOPPFF ADOFIQITER FUNACTURE CEMANENAD VULACUUT FENACOPPFF ADOFIQITER FUNACTURE CEMANENAD VULACUUT FENACOPPFF ADOFIQITER VEAFFF HISSONKOLL NNED VENETERDER VEAFFF HISSONKOLL NNED VENETERDER VEAFFF HISSONKOLL NNED VENETERDER VEAFFF HISSONKOLL NNED NECKKEFFF F	FT	TURN	64	66		
FT HELIX 77 62 FT HELIX 86 96 FT STEAND 107 112 FT STEAND 117 112 FT STEAND 114 122 FT STEAND 114 122 FT STEAND 114 122 FT HELIX 129 136 FT HELIX 141 160 FT HELIX 170 172 TSTEAND 181 183 183 FT HELIX 200 211 FT HELIX 201 234 FT HELIX 203 211 FT HELIX 204 233 FT HELIX 203 307 FT HELIX 303 307 GC MUNALAKKG SADGSVKFL & KAKEDFLKK WEN PONTAH LOPERISTL GTOSFGRVML MUNALAKG SADGSVKFL & KAKEDFLKK WENN FFUNGLIKTST TOSFGRVML <tr< td=""><td>PT HELIX 377 82 PT HELIX 86 96 PT STRAND 107 112 PT STRAND 107 112 PT HELIX 129 136 PT HELIX 129 136 PT HELIX 141 160 PT HELIX 170 172 PT HELIX 203 201 PT HELIX 203 201 PT HELIX 203 201 PT HELIX 244 253 PT HELIX 264 273 PT HELIX 296 298 PT HELIX 296 298 PT HELIX 296 298 PT HELIX 303 307 PT HELIX 346 348 S0 SCUENCE 351 A&: 40600 NW; 59DDD227D2DEEESD CRC64; NCGNALAKKS SQCSVAFLS AKAEDFLKK WEMPAQNTAH LD0FERIKTL GTOSFGRVHL NCGNALAKKS SQCSVAFLS AKAEDFLKK WEMPAQNTAH LD0FERIKTL GTOSFGRVHL NCGNALAKKS SQCSVAFFS HISDDAKAEL NHAD VULACIAL VEMAQNY PFF AD0F10[YEK FUNGTIONED IN SUBJECT OF SU</td><td>FT</td><td>STRAND</td><td>69</td><td>76</td><td></td><td></td></tr<>	PT HELIX 377 82 PT HELIX 86 96 PT STRAND 107 112 PT STRAND 107 112 PT HELIX 129 136 PT HELIX 129 136 PT HELIX 141 160 PT HELIX 170 172 PT HELIX 203 201 PT HELIX 203 201 PT HELIX 203 201 PT HELIX 244 253 PT HELIX 264 273 PT HELIX 296 298 PT HELIX 296 298 PT HELIX 296 298 PT HELIX 303 307 PT HELIX 346 348 S0 SCUENCE 351 A&: 40600 NW; 59DDD227D2DEEESD CRC64; NCGNALAKKS SQCSVAFLS AKAEDFLKK WEMPAQNTAH LD0FERIKTL GTOSFGRVHL NCGNALAKKS SQCSVAFLS AKAEDFLKK WEMPAQNTAH LD0FERIKTL GTOSFGRVHL NCGNALAKKS SQCSVAFFS HISDDAKAEL NHAD VULACIAL VEMAQNY PFF AD0F10[YEK FUNGTIONED IN SUBJECT OF SU	FT	STRAND	69	76		
PT B4 96 PT STEAND 107 112 PT STEAND 114 122 PT HELIX 129 136 PT HELIX 141 160 PT HELIX 141 160 PT HELIX 170 172 PT HELIX 170 173 PT STRAND 173 175 PT STRAND 173 175 PT HELIX 203 205 PT HELIX 208 211 PT HELIX 204 234 PT HELIX 203 307 PT HELIX 346 348 SEQUENCHY HELIX 40620 MY	PT STEAMD 107 112 PT STEAMD 107 112 PT STEAMD 114 122 PT STEAMD 114 122 PT STEAMD 114 122 PT STEAMD 114 122 PT STEAMD 114 126 PT HELIX 129 136 PT HELIX 170 172 PT STEAMD 173 175 PT STEAMD 173 175 PT HELIX 203 205 PT HELIX 203 205 PT HELIX 204 273 PT HELIX 244 253 PT HELIX 244 253 PT HELIX 290 293 PT HELIX 290 293 PT HELIX 296 289 PT HELIX 396 397 PT HELIX 996 293 PT HELIX 996 299 PT HELIX 996 197 PT HELIX 140000 VELOPTELK WENPAONTAM LOOPTELKIL GTOSFGRVML WENFCHER STEALS. 40620 NW; S9DDD227D2DEFESD CPC64; NERALALKEG SEQSSVEFL AKAEUPIEKK WENPAONTAM LOOPTELKIL GTOSFGRVML WENFCHER STEALS. 40620 NW; S9DDD227D2DEFESP FKNNSNLYW NEWVPGOEMF SHLRRLORS EMADEVELA NULLY FYLHS LDLITFOLKP ENLIDQOOF LOUTDFORK RVKGTURE USSGEWEFL ANDUVDUKLY VELAPOPFF ADOPIQIYEK IVSGEWEFS IN SOUCHESD IN RECOKEFSE F	FT	HELIX	77	82		
FT STEARD 107 112 FT STEARD 107 112 FT STEARD 114 122 FT STEARD 114 122 FT STEARD 114 122 FT HELIX 129 136 FT HELIX 141 160 FT STEARD 173 175 FT STEARD 181 183 FT HELIX 203 201 FT HELIX 203 201 FT HELIX 204 233 FT HELIX 290 233 FT HELIX 290 233 FT HELIX 303 307 FT HELIX 304 307 SO SUBJUE 344 34620 SO SUBJUE 344 34620 SO SUBJUE 344 34620 SUBJUE SUBJUE<	FT STEAMD 107 112 FT STEAMD 107 112 FT STEAMD 114 122 FT STEAMD 114 122 FT HELIX 129 136 FT HELIX 141 160 FT HELIX 170 172 FT STEAMD 181 183 FT STEAMD 181 183 FT HELIX 208 211 FT HELIX 208 211 FT HELIX 208 211 FT HELIX 244 273 FT HELIX 244 273 FT TELIX 244 273 FT HELIX 296 295 FT HELIX 296 295 FT HELIX 296 295 FT HELIX 303 307 FT HELIX 346 348 S0 SCUUNCE 351 A&: 40620 HW; 59DDD227D2DEEESD CRC64; NGNAAAKKS S2GSYNEFL & AAREDFLEK WENPAQNTAH LD0FERIKTL GTOSFGRWHL VKHETONHY AKKLDKOKV VKLQIETL MEKRILQAW FFPLVKLES FKNNNLYKV VKHETONHY AKKLDKOKV VKLQIETL MEKRILQAW FFFLVKLES FKNNNLYKV VKHETONHY AKKLDKOKV VKLQIETL MEKRILQAW FFFL HALLYKV VKHETONHY AKKLDKOKV VKLQIETL MEKRILQAW FFFL HALLYKV VKHETONHY AKKLDKOKV VKLQIETL MEKRILQAW FFFL HALLYKV VKHETONHYKV VLAYFYFFL HALLYKV VKHETONHYKV VKLYFYFYFK FFFL HALLYKV VKHETONHYKV VKHETONHYKV VKHETONHYKV VKHETONHYKV VKHETONHYKV VKHETONHYKV VKHETONHYKV VKHETONHYKV VKHETONHYKV VKHETONHYKV VK	FT	HELTY	86	96		
YT STRAND 114 112 YT STRAND 114 112 YT HELIX 121 166 YT HELIX 141 160 YT HELIX 170 172 YT STRAND 173 175 YT STRAND 173 172 YT HELIX 205 211 YT HELIX 244 253 YT HELIX 250 233 YT HELIX 246 348 StOURCHY SAGURANY PERJURATION FOR THE HOURD HAND THANGUNAN HUNDAND H	11 575 575 575 575 575 575 575 575 575 5	FT	STRAND	107	112		
 ISTAND ISTAND	11 DECOND 1:30 1:30 11 DECOND 1:30 1:30 17 HELLX 1:41 1:60 17 HELLX 1:70 1:72 17 STRAND 1:73 1:75 17 STRAND 1:73 1:75 17 HELLX 2:03 2:05 17 HELLX 2:08 2:11 18 DECOND 1:74 HELLX 2:08 2:14 17 HELLX 2:08 2:14 17 HELLX 2:08 2:14 17 HELLX 2:09 2:33 17 HELLX 2:90 2:33 17 HELLX 3:346 3:48 10 SEQUENCE 3:51 AJ: 40:620 HW; 59DDD227D2DEEE5D CRC64; 10 NGAAAKKO SEQUENCEL ANAREDFLKK WEMPAQNTAH LDQFERIKTL GTOSFGKWHL WKNHETONNY AKKLOKGKV VKLKQIEHTL MEKRILQAW FFPLVKLEYS FKNNNLYW WKTVFGGFF SELRADGEK IDMAGENEL GALGYLLY EMAAOFPPFF ADOITDFGFAR ROWGO WTLG GTFERLARET LLSKPWKAY DEALGGLLY EMAAOFPFFF ADOITDFGFAR ROWGO WTLG GTFERLARET LLSKPWKAY DEALGGLLY EMAAOFPFFF	ET.	STRAND	114	122		
1 HELLA 147 140 11 HELA 147 140 11 HELLA 147 140 11 HELLA 147 140 11 HELLA 147 1	1. HELLA 1497 1400 1. HELLA 1497 1400 1. HELLA 1497 1400 1. HELLA 1407 170 1. HELLA 203 105 1. HELLA 203 105 1. HELLA 203 105 1. HELLA 203 105 1. HELLA 203 107 1. HELLA 204 127 1. HELLA 204 127 1. HELLA 206 129 1. HELA 206 129 1.	r I FT	UPLITY	120	126		
11 BELIA 171 100 172 BELIA 171 172 175 BELIA 170 172 175 STRAND 173 175 175 STRAND 173 175 175 STRAND 181 125 175 STRAND 181 125 175 STRAND 181 125 175 STRAND 181 125 175 STRAND 172 175 STRAND 172 175 STRAND 175 175 STRAND 175	1: BELIA 191 192 1: BELIA 194 193 1: BELIA 194 195 1: BELIA 19		UPLIX	129	160		
1 BLLA 1/0 1/2 1 BLLA 1/0 1/2 1 STEAMS 1/3 1/2 1 STEAMS 1/3 1/3 1 BLLA 1/3 1/2 1 STEAMS 1/3 1/2 1 BLLA 203 205 1 HELIX 208 211 1 HELIX 219 234 1 HELIX 244 253 1 HELIX 244 253 1 HELIX 290 293 1 HELIX 290 293 1 HELIX 290 293 1 HELIX 303 307 1 HELIX 303 307 1 BLIX 303 307 1 BLIX 303 307 1 BLIX 303 307 1 BLIX 303 307	11 BELIA 1/0 1/2 17 STRAMD 173 175 17 STRAMD 173 175 17 STRAMD 131 105 17 STRAMD 131 105 17 STRAMD 131 105 10 STRAMS SQUENCE 351 AJ: 40 SQUENCE 350 SQUENCE 351 AJ: 40 SQUENCE 350 SQ		NEL IX	141	100		
1 STRAND 1/3 1/3 1 BLIX 203 205 1 BLIX 200 205 1 BLIX 200 201 1 BLIX 201 201 1 BLIX 201 201 1 BLIX 203 201 1 BLIX 204 203 1 TUEN 286 283 1 HELIX 206 293 1 HELIX 303 307 1 HELIX 346 348 3 SQUENCY SSLOESOKER SELSEN COLON MY SSDUENCY 1 WTHEREN AND FULKERER FINNENARY HUTTE CONTRAL FULK HUTENDER FINNENARY 1 WENDERSON VERENDERSON VERENDERSON FINNENARY	FI ADP 143 TAND 143 143 TSTAND 181 183 TT HELIX 203 205 TT HELIX 203 205 TT HELIX 203 205 TT HELIX 208 211 TT HELIX 214 233 TT HELIX 244 233 TT HELIX 296 293 TT HELIX 296 293 TT HELIX 396 307 TT HELIX 396 307 TT HELIX 396 307 TT HELIX 396 307 TT HELIX 303 307 TT HELIX 196 298 TT HELIX 196 29	r I FT	OTDAND	170	172		
1 SirAnu 101 103 1 SirAnu 101 103 1 BLIX 203 214 1 SirAnu 203 214 1 BLIX 203 214 1 BLIX 204 214 1 HELIX 244 253 1 HELIX 264 273 1 HELIX 290 293 1 HELIX 303 307 1 HELIX 344 34500 MF 20 SSIALALAKG 350SSVKFL KAKEDFLKK 30 307 HELIX 344 30 307 HELIX 34500 MF 20 SSIALALAKG 350SSVKFL KAKEDFLKK HENONTH 30 SOSIALAKG 350SSVKFL KAKEDFLKK HENONTH 20 SSIALALAKG 350SSVKFL KAKEDFLKK HENONTH 30 SOSIALAKGO 350SSVKFL KAKEDFLKK HENONTH	1 SIFANU 101 103 1 SIFANU 101 103 1 KELIX 203 211 1 KELIX 203 211 1 KELIX 203 211 1 KELIX 204 253 1 THELIX 244 253 1 THELIX 244 253 1 THELIX 244 253 1 THELIX 296 298 1 HELIX 296 298 1 HELIX 303 307 1 HELIX 304 348 30 SEQUENCE 351 A&: 40620 HW; 59DDD227D2DEEE5D CRC64; 1 KGNAAAKKS SEGEWKFL AKAEDFLKK WENPAQNTAH LD0FERIKTL GTOSFGRVHL 1 KGNAAAKKS SEGEWKFL AKAEDFLKK WENPAQNTAH LD0FERIKTL TOFOFFF 1 GOTFFGFK SUNGTTUC GTPERIAND VIALOUT VIALOUTES FKNMANTAW 1 GOTFFGFK SUNGTTUC GTPERIAND VIALOUTEN FKGNLXNOWN DINNHKWFAT 1 DUHAIYORK VEAFFFFK GFGDTSNFDD YEEEEIRVSI NEKCGKEFSE F	2.1	STRAND	173	175		
FT HELIX 203 205 FT HELIX 208 211 FT HELIX 208 211 FT HELIX 219 234 FT HELIX 244 253 FT HELIX 244 253 FT HELIX 246 278 FT HELIX 296 288 FT HELIX 303 307 FT HELIX 306 348 S SQUENCE 351 A4: 40520 HW: S9DDD227D2DEEESD CRC64: MCNAAAKKG SDCSVKEFL AKAKEDFLKK WENPANTAH LOVERIKTI GTOSFGRVHL MCNAAAKKG SDCSVKEFL AKAKEDFLKK WENPANTAH DIVINKE FINLIDQOGY IQVTDFGAK RVKGTUTIC GTPETLAPINATUUG TILIKKYTAK DWUALOUTI FINLAGYPFF AUGPIGITEK VSGKWEFTS HISSOKUML KNUMLYNDIK FFGULNGVA DINNKKFFAT TUTIAITUGK VEAPFIPHER GOCTISTED YEEEGINGS NEKCGEFS F	TT HELIX 203 205 TT HELIX 203 205 TT HELIX 208 211 TT HELIX 219 234 TT HELIX 244 253 TT HELIX 264 273 TT TURN 286 263 TT HELIX 303 307 TT HELIX 303 307 TT HELIX 346 348 OS SEQUENCE 351 A4. 40620 NW; S9DDD227D2DEEE5D CPC64; NGRAAAKKG SEQESVKEFL AKAREDFLEK WENPAONTAM LDQFERIKTL GTOSFGRVML NUMERATED SUCESVKEFL AKAREDFLE NNEL NUMERATED SUCESVKEFL SUCESVKEFL AKAREDFLE NNEL NUMERATED SUCESVKEFL AKAREDFLE NNEL NUMERATED SUCESVKEFL AKAREDFLE NNEL NUMERATED SUCESVKEFL SUCESVKEFL SUCESVKEFL NUMERATED SUCESVKEFL SUCESVKEFL SUCESVKEFL NUMERATED SUCESVKEFL SUCESVKEFL SUCESVKEFL NUMERATED SUCESVKEFL SUCESVKEFL SUCESVKEFL NUMERATED SUCESVKEFL SUCESVKEFL NUMERATED SUCESVKEFL SUCESVKEFL NUMERATED SUCESVKEFL SUCESVKEFL NUMERATED SUCE	r T	STRAND	181	183		
FI HELIX 205 211 FI HELIX 219 234 FT HELIX 244 253 FT HELIX 264 273 FT TURN 266 269 FT HELIX 290 293 FT HELIX 290 293 FT HELIX 303 307 FT HELIX 303 307 FT HELIX 346 348 S SEQUENCE 351 A4: 40520 HW; S9DDD227D2DEEESD CRC64; NGNAAAAKG SEQESVKEFL AKAKEDFLKK WENPAQNTAH LOVERIKTI GTGSFGRVHL VKHHETGNEV VKKNEFTL NEKRICAAN PFU/KLEFF FKNNSNLYHV NEVPFOGENF SHLBALGKE DEMARKALO ULL FTLBS LDLIYPDLKF FNLSDQOF IQUTDFGAK RVKGTURIC GTFEYLDAMERKALO ULL FTLBS LDLIYPDLKF FNLSDQOF ADOPIQITYEK IVSGKWHTS HISSOLKANDL KHAW_VDWIALGULY HAAGYPFFF ADOPIQITYEK IVSGKWHTS HISSOLKANDL KHAW_VDWIALGULY HAAGYPFFF ADOPIQITYEK IVSGKWHTS HISSOLKANDL KHAW_VDWIALGULY HAAGYPFFF	ri HeLiX 200 211 FI HELIX 219 234 FI HELIX 219 234 FI HELIX 244 253 FI HELIX 244 253 FI HELIX 296 253 FI HELIX 296 253 FI HELIX 303 307 FI HELIX 303 307 FI HELIX 346 348 S0 SEQUENCE 351 A&: 40620 HW; 59DDD227D2DEEESD CRC64; HOGNAAAKKO SPCSTVERFL ARAMEDFLKK WENPAONTAM LDOFERIKTL GTOSFGKVML VKNHETONHY AKKIDSCOV VELKQIEHTL NEKRILQVM FPFLVKLES FKNNNLYW HETVPGCHF SILKAJCKE ENABEDFLKK WENPAONTAM LDOFERIKTL GTOSFGKVML VKNHETONHY AKKIDSCOV VELKQIEHTL NEKRILQVM FPFLVKLES FKNNNLYW HETVPGCHF SILKAJCKE ENABEDFLKK WENPAONTAM LDOFERIKTL GTOSFGKVML VKNHETONHY AKKIDSCOV VELKQIEHTL NEKRILQVM FPFLVKLES FKNNNLYW HETVPGCHF SILKAJCK CONSTRUCTION FFALARET LISKOVAK FFALAGULAVI FEAAGYPPFF ACTION FFALARET FILSKOV VELKGIEHTL NEKRILQVM FPFLVKLES FKNNNLYW HETVPGCHF SILKAJCK CONSTRUCTION FFALARET LISKOVAK FFALAGULAVI FEAAGYPPFF ACTION FFALARET FILSKOVAK VELKGIEHTL NEKRIKVAX TDVILAIYQEK VEAPFIPKFK GFGDTSNFDD YEEELIEVSI NEKCGKEFSE F	t T	HELIX	203	205		
FT HELIX 219 234 FT HELIX 244 253 FT HELIX 244 253 FT HELIX 244 253 FT HELIX 264 273 FT HELIX 250 253 FT HELIX 303 207 FT HELIX 303 207 FT HELIX 346 348 O SEQUENCE 351 A4: 40620 NV: S9DDD227D2DEEESD CRC64: NGNAAAKKG SEQUENCKFL KAKEDFLKK WENPAONTAH LOPERIKTL GTOSFGRVHL NGNAAAKKG SEQUENCKFL KAKEDFLKK WENPAONTAH LOPERIKTL GTOSFGRVHL NGNAAKKG SEQUENCKFL KAKEDFLKK WENPAONTAH LOPERIKTL GTOSFGRVHL NGNAAKKG SEQUENCKFL KAKEDFLKK WENPAONTAH DINGKFL NGNAAKKG SEQUENCKFL KAKEDFLKK WENPAONTAH LOPERIKTL GTOSFGRVHL NGNAAKKG SEQUENCKFL KAKEDFLKK WENPAONTAH LOPERIKTL GTOSFGRVHL NGNAAKG SEQUENCKFL KAKEDFLKK WENPAONTAH LOPERIKTL GTOSFGRVHL NGNAAKKG SEQUENCKFL KAKEDFLKK WENPAONTAH LOPERIKTL GTOSFGRVHL NGNAAKKG SEQUENCKFL KAKEDFLKK WENPAONTAH LOPERIKTL GTOSFGRVHL NGNAAKKG SEQUENCKFL KAKEDFLKK WENPAONTAH FUNKLESKER AUDPLOTYFK VYSGNAATS HISSOLKALL KANULYNDIK SEGGREFSE F	FT HELIX 219 234 FT HELIX 244 253 FT HELIX 264 273 FT HULX 264 273 FT HULX 290 293 FT HULX 290 293 FT HULX 296 289 FT HULX 396 300 SO REQUENCE 351 AA. 40620 NW; S9DDD227DIDEEE5D CRC64; KENALAKKO SECSVEFL AKAEDFLEK WENPAONTAM LDQFERITL GTOSFGRVML VUENETONHY ANKLDROW VULKUETH. NERFUNANTAM FULVELEES FKNISNLYNV NEWFVFGGEN SHLRBIGES ENALEVALO ULL FFYLHS LDLIYRDLKP ENLLDQGOY IQVTDFGAR RVKGTURLG CFTPLIARFNIL HNED VDLIYFPL IVSG AVEFES HI SJOLKDEL NHED ADOPIGIYEK IVSG AVEFES HI SJOLKDEL NHED VELEFTURG FFS HI SJOLKDEL HNED VELEFTURG FFS HI SJOLKDEL NHED VELEFTURG FFS HING FFS HI SJOLKDEL NHED VELEFTURG FFS HING FFS HING FFS HING FFS HING FFS FKNING ADOPIGIYEK VELEFTURG FFS FFS FF FFS HING FFS HING FFS HING FFS HING FFS FKNING FFS FKNING FFS FKNING FFS FFS FFS FFS FFS FFS FFS FFS FFS FF	rΤ	HELIX	208	211		
FT HELIX 244 253 FT HELIX 264 273 FT HELIX 264 273 FT HELIX 290 283 FT HELIX 296 298 FT HELIX 303 307 FT HELIX 346 348 Stoutence 351 A4: 40620 NW; Stoutence 351 A4: 40620 NW; VKINETCREV Stoutence Stoutence Stoutence INFUNCTIONAR WKKSDERDERS Stoutence Stoutence Stoutence INFUNCTIONAR WKKSDERDERS Stoutence Stoutence Stoutence INFUNCTOR WARTENDERS Stoutence Stoutence INFUNCTOR WARTENDERS Stoutence Stoutence INFUNCTOR VEXENTERS INSCOMENTS INSCOMENTS INFUNCTOR VEXENTES INFUNCTOR INFUNCTOR ANDOPIOTYEK VSORTHER GODTSTOND VEXENTERSON INECOCRESE F INFUNCTOR	FT HELIX 244 253 FT HELIX 244 253 FT HELIX 246 273 FT TURN 286 289 FT HELIX 290 293 FT HELIX 303 307 FT HELIX 303 307 FT HELIX 304 346 S0 SEQUENCE 351 A&: 40620 HW; 59DDD227D2DEEE5D CRC64; NGNAAAAKKO SEQESVKEFL AKAKEDFLKK WENPAONTAM LDQFERIKTL GTOSFGRVML VKHRETONNY AMKLDROKY VEKADETH LEKKTLAVAN DWALGUTIY ETHACF9FF NGNAAAKKO SEQESVKEFL AKAKEDFLKK WENPAONTAM LDQFERIKTL GTOSFGRVML VKHRETONY AMKLDROKY VEKADETH LEKKTLAVAN DWALGUTIY ETHACF9FF ADQFIQIYEK IVSGKWAFFS HISSDEKEL NNEL VECKEFSE F	rТ	HELIX	219	234		
FT HELIX 264 273 FT TUEN 266 269 FT HELIX 290 293 FT HELIX 290 293 FT HELIX 303 302 SEQUENCE 313 43 4020 MSNAAAKG SEQUENCE 314 40420 WHIETCHNY MKILDEGOKY VKINETCHNY MKILDEGOKY VKINETCHNY MKILDEGOKY VKINETCHNY MKILDEGOKY JQUTDFGAK KVKG VKINETCHNY MKILDEGOKY VKINETCHNY JQUTDFGAK KVKGATUTIC CTEVILLEK VKINETCHNY VKINETCHNY JQUTDFGAK KVKGATUTIC TISZNIKAVA VKINETCHNY VKINETCHNY JQUTDFFKK VSGAVKFFS HISSNIKAVA VWINKUFAT TUTAINUGK VKINETCHNY VKINETCHNY VKINETCHNY JQUTDFFK VSGAVKFFS HISSNIKAVA VWINKUFAT TUTAINUGK VKINETCHNY VKINETCHNY VKINETCHNY JQUTDFFK VSGAVKFFS HISSNIKAVA VKINKUFA	FT HELIX 264 273 FT TURN 286 289 FT HELIX 290 283 FT HELIX 290 283 FT HELIX 303 307 FT HELIX 304 307 SUBJECT BELIX 34620 NW; SUBDR27D2DEEED CRC64; SUBJECT HELIX 34620 NW; SUBDR27D2DEEED CRC64; SUBJECT HELIX 34620 NW; SUBDR27D2DEEEED CRC64; SUBJECT HELIX 34620 NW; SUBJECT HERACONAL EDUFFERITI GTOSFGRVML WENNETONY ANKIDENCY VELADETH INFRINAVA UPWIAGULTY ETHACTOPFF AUGPTOINTS HISDERSE HEMARYANG HIL FFVHS LDLITEDLEF ENLIDQOOY HEVPFOCHT SHIEPLORE EMARYANG HIL FFVHS LDLITEDLEF ENLIDQOOY HOUTOFAR RVKGTHIC GTPEYLARD HILFT LISSTMAND UPWIAGULTY ETHACTOPFF AUGPTOINTER INSOLWEFS HISDERDELE NHEL VENETOR VELAPFIPKFK GPGDTSNFDD YEEEEIRVSI NEKCGKEFSE F	FT	HELIX	244	253		
<pre>FT TURN 286 289 FT HELIX 296 293 FT HELIX 296 293 FT HELIX 296 293 FT HELIX 303 307 FT HELIX 3103 307 SSQUENCE 351 A&: 40620 NW; S9DDD227D2DEEE5D CRC64; NGNAAAAKKG SEQESVKEFL AKAKEDFLKK WENPAQNTAH LOFERIKTL GTOSFGRVHL VKHETGRVKV VKLKGIETH. NEKRIGAAN PFUVKLEFFF FKNNSNLYHV HEYVPGGHF SHLRADCGEF URUAFFLAG TULFFTHS LDITYDLKP FNLLTDQGF IQVTDFGAK KVKG TUTIG GTFETLAGET LISKYVNAV DUWLAUCULY ENLACYPFFF ADQF10TFEK IVSGKVEFF GFG7SNENDY HEKCGKEFS FF </pre>	<pre>FT TURN 286 289 FT HELIX 290 293 FT HELIX 290 293 ST HELIX 303 307 FT HELIX 303 307 FT HELIX 3046 346 S SEQUENCE 351 A4: 40620 NW; 59DDD227D2DEEE5D CRC64; NGNAAAAKKG SEQESVKEFL AKAKEDFLKK WENPAQNTAM LDQFERIKTI GTGSFGKVML VKHETGNVT AKKLDFGKE DEHADEVAG UUL FFLHS LDLIVRDLKF ENLLIDQQGY IQUTDFGAR RVKGTUTUC GTFERILARD LINKUUT FFLHS LDLIVRDLKF ENLLIDQQGY IQUTDFGAR RVKGTUTUC GTFERILARD LINKUUT FFLHS LDLIVRDLKF ENLLIDQQGY ADQFIQIYEK IVSGKWAFFS HISSDEKEL NHELUVDITK RFGNLKNGVN DIKNKWFAT TDUIAIYQRK VEAPFIPKFK GFGDTSNFDD VEEEEIRVSI NECGKEFSE F</pre>	FТ	HELIX	264	273		
<pre>FT HELIX 290 293 FT HELIX 296 296 FT HELIX 296 296 FT HELIX 305 307 FT HELIX 304 307 FT HELIX 304 307 FT HELIX 304 307 FT HELIX 304 500 HY; S9DDD227D2DEEESD CRC64; S0 SCHALAKKG SPGSYKEFL KKKEDFLKK HENPAGNTAH LOFTERIKT. GTGSYGRVML VKHETGHOVY VKKIGTERIKT NEKRIGAAN DFWLVKLEFF FKNNSNLYHV HEVVPGGENF SHLEDCOFF ERMANFLAG. THI FFTHES LDLIYPDLKF FNLLDOGOY IQUTDFGAK KVKGITURIG GTFEYLAPELI LISKYTMAKA DWALGULIY ENLACYPFFF ADQFIGITEK IVSGNTHFFF GFGJSNLHDL HHMLVDLTK FFGHLKNOWN DINNKKFFAT TDWIAIVGK VEAFFFFFF GGGJSNLHDL HHMLVDLTK FFGHLKNOWN DINNKKFFAT TDWIAIVGK VEAFFFFFF GGGJSNLHDL HHMLVDLTK FFGHLKNOWN DINNKKFFAT </pre>	FT HELIX 290 293 FT HELIX 290 296 FT HELIX 303 307 FT HELIX 304 346 S0 SKOUENCE 351 AJ: 40620 NW; 59DDD227D2DEEESD CRC64; NGNAAAKKG SQCSWEFL AKAEDFLKK WENPAGNTAH LDOFERIKTL GTSTGRVHL NGNAAAKKG SQCSWEFL AKAEDFLKK WENPAGNTAH LDOFERIKTL GTSTGRVHL NGNAAAKG SQCSWEFL AKAEDFLKK WENPAGNTAH LDOFERIKTL GTSTGRVHL NGNAAAKG SQCSWEFL AKAEDFLKK WENPAGNTAH LDOFERIKTL GTSTGRVHL NGNAAAKG SQCSWEFTE HENDIG LOUTDFGAR RUKGTUTU CETPELIAND ADOFLOIYEK IVSGN VAFFS HESSEKKEL NHEL VDLTK FFGHLKNGVN DIKNHKWFAT TDWIAIYORK VEAPFIPKFK GPGDTSNFDD YEEEEIRVSI NEKCGKEFSE F	FT	TURN	286	289		
FT HELIX 296 298 FT HELIX 303 307 FT HELIX 303 307 FT HELIX 304 348 S SEQUENCE 351 A&: 40620 NW; S9DDD227D2DEEE5D CRC64; MONAAAAKKG SEQESVKEFL AKAKEDFLKK WENPAQNTAH LOOFERIKTL GTGSFGRVHL VKHEETGNNY AMKILDKOKV VKLKQIEHTI NEKRILGAND FPUVKLEFS FKNNSNLYHV HEYVPGGENF SHLBADCESE ENADEYLAD, TULFFULHS LDLIYADLKP ENLLIDQQGY IQVTDFGAK KVKGNTUTIG GTFEVLADET LISKYTVAKV DUWALGVLIY ENLAGYPFFF ADQF10TEK IVSGNTHFF GF07SNLKDL KNEWLVDIX FFGULKNOVM DINNKKFFAT TDVIAIVGK VEAFFFFKF GF07SNLKDL KNEWLVDIX NEKCGKEFSF F	FT HELIX 296 296 FT HELIX 303 307 FT HELIX 303 307 FT HELIX 346 346 S SEQUENCE 351 A4: 40620 NW; S9DDD227D2DEEESD CRC64; MGNAAAAKKO SEQUENCKV VEKAGENTL DEKRIGAOWAN FPLVKERFS FKNSNLYNV NEWUPGOEHF SHLARIOFS EDHAERYAG INI FEVLHS LDLIVEDLKP ENLLDQQOY IQUTDFOAR RVKGTUTUL GTPEYLAPEI LISKPINAAD WWALGULY ENAACPPFF ADOPIQIYEK IVSG KWAFS HISSDERDE NHEL VULT REACHEFSE F	FΤ	HELIX	290	293		
FT HELIX 303 307 FT HELIX 346 348 S0 SEQUENCE 351 AA; 40620 MW, 59DDD227D2DEEE5D CRC64; NGNAAAKKG SPOSYMERFL AKAKEDFLKK WENPACNTAH LDOFERIKTL GTOSFGRVHL VKHMETGNHY AMKILDKOKY VKLKOIEHTL NEKRILQAVN PPFLVKLEFS FKINSNLYMY MEYVPGGENF SHLIPACOKE DEMOFIAL HILTEYLHS LDJUTYEDLKP ENLLIDQOOY IQVTDFGAK KVKGMINTLG GTEETLAPEI ILSKUTKAV DUMALOULT ENLACIPPFF ADQF10TEK IVSGKWAF5 HISSUNKAL KNELVUTLK KFGMLENNOVA DINNKKFAT TUTAIATUGK VEAFFIFKR GPOTISMED VEEEEIKVES INEKCGKEFSE F	FT HELIX 303 307 FT HELIX 346 348 S0 SEQUENCE 351 AX; 40620 NW; 59DDD227D2DEEE5D CRC64; NGNAAAKKG SEQSEVEFL AKAEMPIKK WEMPAQNTAH LDQFERIKTL GTOSFGKVML VKNHETONNY AKKLDKOKV VKLKQIEHTL MEKRILQAVM FPFLVKLEFS FKNNSNLYRV HETVFGGKTF SELAFGACE CHPEYLVKLTILLGFTVLKHE PLALLOGOFF ADOPIQIYEK IVSGN VNFFS HISSDKKEL NHER VVLLTK FFGNLXNONN DIRNHKWFAT TDWIAIYQRK VEAPFIPKFK GPGDTSNFDD YEEEEIRVSI NEKCGKEFSE F	FΤ	HELIX	296	298		
FT HELIX 346 348 ST HELIX 346 348 ST SEQUENCE 351 AJ: 40620 NW; S9DDD227D2DEEESD CRC64; MCNAAAAKKG SEQESVKEFL AKAKEDFLKK WENPAQNTAH LOPERIKTL GTOSFGRVML VKHETCHNY AMKLIDAGVK VKLKQIEHTI NEKRILGANN PFUYKLEFS FKNNSNLYMV MEYVPGGENF SHLBDICGES ENADEYLAD, TULFFUHS LDLIYADLKP ENLLIDQQGY IQUTDFGAK KVKGTUTLG GTPEYLADEI ILSKYTWAYA DUWALOULY ENLAGYPFFF ADQPIQITEK IVSGNWEFFS HIJSSDLKDL KNEWLVDLK FKGLKNOVM DINNKKFFAT TDWIAIVGK VEAFFFFKF GOPOTSMED VEEEGIKENS INEKCGKERSF F	FT HELIX 346 348 ST SEQUENCE 351 A&: 40620 NW; 59DDD227D2DEEE5D CRC64; MCNAAAKKO SEQESYKEFL AKAKEDFLKK WENPAQNTAH LDQFERIKTL GTOSFGRVHL VKHHETONY ANKIDAKOKV VKLOEITI LEKKIDAVAN FPLVKHESF SKUNSNLYW MEYVPGCEMF SHLRBICHES EDHAGEYLAG IULTFEILS IQVIDFORA RVKGTWITLG CTPERILADE ILSKÖWAKAV DWALGULTY ELMACPPFF ADQFIQIYEK IVSCHWAFFS HISSDEKDE NHEL VDUTALYQRK VEAPFIPKFK GPGDTSNFDD YEEEEIRVSI NERCGKEFSE F	FΤ	HELIX	303	307		
SO SEQUENCE 551 AJ; 40620 NW; S9DDD227D2DEEESD CRC64; NGNAAAKKG SSOSSWEPEL AKKEDELKK WEDNAONTAL LOGGENEIL, GTOSFGRVHL VEHHETONHY AMKILDKOKV VELKOIEHTL NEERILGANN PFELVKLEFS FENNENLYHV HEYVPGGENF SHLER <u>LGERE EBMADFWAG</u> INLEFETLAD. INLEFETLAD. IQUTDFGK KVKGHTULG GTETLAPEI ILSK VIKAV DUNALOULIY EHAACYPFF ADQF10ITEK IVSG KVKT5 HISSDENKL KNEUV UDLTK FFGMLENNOVN DINNKFFAT TUTAIJOKK VEAFFFFKF GPCJTSWEDV HEEKGKEFSE F	SQ SEQUENCE 351 AA: 40620 NW: 59DD227D2BEESD CRC64: NGNAAAKKG SEQSWERFL AKAEPGIKKE WERAGANTAH LOPFENTITA GTOSFGRUHL VENHETONNY AMKILDRORV VELKQIEHTL DEKRILQAWI FPFLVKLEFS FKNONNLYNV NEVVFGGEFF SHLPAGERS EMALEFALAD. HLTFFVHAR ENLLTQOGGY IQUTDFGFAR RVKGTUTLC GTPFYLAPEI ILSKDTNKAV DWWALGVLIY EMAAGYPFFF ADQPIQIYEK IVSGKVRF5 HISSDLKDE ANELOVDLTK RFONLKNGVD DIANKWFAT TDWIAIYQRK VEAPFIPKFK GFGDTSNFDD YEEEEIRVSI NECCKEFSE F	FΤ	HELIX	346	348		
NGNAAAAKG SEQESYKEFL AKAKEDFLKK WENPAQNTAH LOQERIKTL GTOSFGWHL VKHHETGNNY ANKLIDAGKV VKLKQIEHTI NEKRILQAN PFUYKLEFS FKNNSNLYRV HEYVPGGEHF SHLBR <u>LGEST EFRANFYLAG</u> ILLFFYLHS LDLIYRDLKP ENLLIDQGGY IQUTDFGAK RVKGITUTIG GTPEYLAPEI LISKYTNKAV DWALQULY FRAACYPFFF ADQFIQIYEK IVSG ITUTIG GTPEYLAPEI LISKYTNKAV DWALQULY FRAACYPFFF IDUTAIYORK VEAPFIPKER GOPOTISHED YEELEKING NEKCGKEPSE F	NGNAAAAKG SEQESVKEFL AKAKEDFLKK WENPAQNTAH LOGFERIKTL GTOSSGEVHL VKHRETONHY ANKLORGOV VKLOGIENTL NEKRIAOVAN FPLIVKEFS FKDNSNIJWV NEVVPGCENF SHLBRLORE EPHAEPYIAO IULTFEILSK DIVIGUAL PERACHOPFF AUGPIQITEK IVSG INTELG CITELIUSTEL INKLU VDLTK FGNLKNGVN DIXNNKWFAT TDWIAIYQRK VEAPFIPKFK GPGDTSNFDD YEEEEIRVSI NEKCGKEFSE F	SQ	SEQUENCE	351 AA	; 40620	NW; 59DDD227D2DEEE5D CRC64;	
VKHHETONHY AMKILDKOKU VKLKOIEHTL NEKRILOANN PFÜLVKLEFS FKONSNLYHV NEVVGGENF SKLEPAGGES EDMADEKAL INLEFELIS LOLIYODKUE ENLLIDOOGY IQVTDFGFAK KVKO <mark>HTUTLC GTPEYLAPEI ILSKO</mark> YNKAV DUWALGVLIY ENAAGTPFF ADQFIGITEK IVSG KVKIFS HISSOLKUL KNUL VULTK RFGMLENOVU DINNKKFAT TUTAILYGK VEAFFIFKK GODISTRUD VEEDELKKUS NEKCGKERSE F	VKHRETCNHY AHKILDRORV VELKQIEHTL NERKEILQAVE FPFLVRLEFS FRONSNLYHV NEYVGCHF SHLBALENSE EMALEVIAL ILLFFYLIAE LINEROVER ENLILGOGGY IQYTDFGFAR RVRGTUTLC GTPFYLAPEI ILSKONKAV DUWALGVLIY ENAGYPPFF ADQPIQIYEK IVSGV AFFS HISSDERDE NHEL WULTK RFONLKNGAN DIRNKWFAT TDWIAIYQRK VEAPFIPKYK GFGDTSNFDD YEEEEIRVSI NECCKEFSE F		MGNAAAAKKG	SEQESV	KEFL AKAI	KEDFLKK WENPAONTAH LDOFERIKIL GIGSFGRVML	
NETVPSCENF SULER <u>JCEFF SUNAFFIAG TVITF</u> TINS IDLITNDLKP ENLLIDQOGY IQUTDFCRK RVKGITULG CTFFILJELI LISKYTNKAV DUWALGULI VENLAGYPFF ADGPIGITEK IVSG IVETS HISSUNDL RNDLU VDLTK RFGNLENGVA DINNKKFFAT TDUTAIVGRK VEAFFIFRFK GPGJTSHDD YEEEEIRVSI NEKCGEFFS F	METVPGCEMF SHUR <u>PLORES EDNATEVING TUL T</u> FYIHS LDLIYEDLKP ENLLIDGOGY IQUTDFGAR RVKGTUTLC GTPFYLAPFI LISR/NYKAN DUWLACULT ENLACYDFFF ADGPIGIYEK IVSG WAFFS HISSDEKDLE NHED UVDLTK RFGNLKNGVM DINNHKWFAT TDWIAIYORK VEAPFIPKFK GPGDTSNFDD YEEEEIRVSI NEKCGKEFSE F		VKHMETGNHY	AMKILD	KOKV VKLI	QIEHTL NEKRILQAVN FPFLVKLEFS FKDNSNLYMV	
IQUTDFGFAR RVKG <mark>H</mark> TWILC GTFEYLAPEI ILS <mark>KD</mark> YNKAV DWALGULIY ENAAGYPFFF Adqpigitek IVSG NRTJ HISSDIRNGL RNHG VULIK FRAGNANOVN DIRNKFFAT TDWIAIVGR VEAFFFRFK GPOTISNED YEEDEINSI NEKCGREPSE F	IQUTDFGFAR RVKGTUTLC GTPEYLAPEI ILSKDYNKAV DUBALGVLIY ENALGYDPFF ADQPIQIYER IVSGTV AFFS HISSDERDEL NNEL GVDLTK RFGNLKNGVN DIRHKWFAT TDWIAIYQRK VEAPFIPKYK GPGDTSNFDD YEEEEIRVSI NEKCOKEFSE F		MEYVPGGEMF	SHLRRI	CDFS FDW	DEVAAO JULTFEYLHS LDLIYRDLKP ENLLIDOOGY	
ADOPIQIYEK IVSG AVREFS HESSELKUL KHEL VULIK REGNLKNOVM DIKNHKWFAT TDWIAIYQRK VEAPFIPKEK GPGDTSNEDD YEEEEIRVSI NEKCGKEFSE F	ADOPIQIYEE IVSO NAFFS HISSDERDEL NULD UDLTE PFORLENGUN DINNHEWPAT TDWIAIYORE VEAPFIPEFE GPGDTSNFDD YEEEEIRUSI NEKCGKEFSE F		IOVTDEGEAR	RVKGET	WTLC GTPI	VLAPET ILSK VNKAV DHWALGULTY EMAAGYPPFF	
TDUIAIYQRK VEAPFIPKFK GPGDTSNFDD YEEELIRVSI NEKCGKEFSE F	TDWIAIYQRK VEAPFIPKFK GPGDTSNFDD YEEEEIRVSI NEKCGKEFSE F		ADOPTOTYER	TVSG		DVDLTK REGNLKNGVN DIKNHKNEAT	
IPPEREIQUE TERITETRIE OF OPTIME PP IEBEETETET HEROOREISE I			TDUTATYORN	VEADET	PKEK GPGI	TENERD VEFETURET NEVCOVERES F	
			QRF	VEARIT	. AFK OFGI	INTER TREBERING AND AND A TREASTREE OF I	
				_			2

According to Swissprot, both T1 and T3 are possible phosphorylation sites. If you really needed to know which was the case here, or whether it was a mixture, you'd have to acquire more data. Maybe try a different enzyme or target the incomplete cleavage peptide that includes the preceding KG so as to move the sites towards the centre of the peptide, where you might get stronger b and y fragments



If you are using Mascot 2.3 or earlier, the delta score calculation is not performed in Peptide View. These are our suggested guidelines when using Mascot for site analysis:

If alternative sites differ by 20 in score, safe-ish to disregard lower one(s)

If alternative sites have similar scores, you may be able to choose one by inspection. But, be careful ... one peak is just one peak

Often, you just can't differentiate between adjacent sites, even with great data.



Now, back to the challenge of finding PT modifications. There are many hundreds of modifications in Unimod, yet I've emphasised the importance of using the minimum number of variable modifications in a search. So, how are we supposed to find unusual modifications?

If you are searching uninterpreted MS/MS data, the efficient way to find unusual modifications, as well as variations in the primary sequence, is a two pass search. The first pass search is a simple search of the entire database with minimal modifications. The protein hits found in the first pass search are then selected for an exhaustive second pass search. During this second pass search, we can look for all possible modifications, sequence variants, and non-specific cleavage products.

Because only a handful of entries are being searched, search time is not an issue. It would be extremely difficult to calculate meaningful statistics for the additional matches in an error tolerant search, and we don't report expect values. The evidence for the presence of any particular protein are the matches from the first pass search. The additional matches from the second pass search serve to increase coverage and may discover interesting modifications or SNPs.



For modifications, an error tolerant search looks for one unsuspected modification per peptide in addition to those mods specified as fixed or variable. This is sufficient because it will be rare to get two unsuspected mods on a single peptide

Error Tolerant Search	
Primary sequence variants	
Protein database	
Look for all residue substitutions	
No attempt to identify single base insertions deletions because of frame shifts	£
 Nucleic acid database 	
Look for all single base substitutions, insertions & deletions	ons
MASCOT : Modifications © 2007-2012 Matrix Science	MATRIX

The error tolerant search also looks for sequence variants, such as single nucleotide polymorphisms (SNPs) or sequencing errors.

For a protein database, we can't look for the consequences of inserted or deleted bases, because these give rise to frame shifts, and the entire sequence changes from that point on.



There are some constraints on the standard, first pass search



Otherwise, submitting the search is just like submitting a standard search except that you check the Error Tolerant Checkbox



You see two sets of progress reports

🗐 Pept	tide	Summa	ary Report (Er	ror tolerant e	xample) - Micro	osoft Intern	et Exp	lorer							
File	Edit	View	Favorites To	iols Help							and had been a state of the sta				
G	Back	- 6) - 💌 💈) 🟠 🔎 s	earch $ \lacelow \l$	rites 🚱	8.	2	🔁 -	Pot	vermarks 🌃 \land %				
: Addres	- A	http://s	www.matrixscienc	e.com/cni/master	results.pl?file=la	lata/20070626	FoGeri	e5.dat		1	V 🗗 60				
-															
Sel	ect A		Select None	Searc	ch Selected	Ентог	tolera	ant							
1.	<u>×</u>	AAS17	08 Mass	: 56371 !	Score: 782	Queries	s mat	ched:	27 emPl	AI: 0	.78				
	Check to include this hit in error tolerant search														
	Dury Observed Mr/emb) Mr/oplo). Dolta Mise Sone Evment Dank Dantide														
	Qu	iery	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Expect	Rank	Peptide				
6	~	27	462.6807	923.3468	923.5116	-0.1649	0	33	16	1	R.FPYVALSK.T				
		41	517.1760	1032.3375	1032.5604	-0.2229	0	70	0.0036	3	R.GSSIFGLAPGK.A				
		62	564.6804	1127.3463	1127.5764	-0.2301	0	10	2.8e+03	6	R.GFFLFVEGGR.I				
	~	65	567.6567	1133.2987	1133.5499	-0.2511	U	44	1.1	1	K. GREVISVARK.A + Uxidation (A)				
		100	653 2101	1226.3836	1226.6329	-0.2473		20	5 70-05	2	K.LOPEIPLANDK.F + OXIGATION (H)				
		124	710.2235	1418 4324	1418 7154	-0.2829		95	5.70-05	÷.	K CNFOTTCISABAB F + Acety] (N-term): $f+72$ 0211 at N-term Cl				
		12.6	726.1806	1450.3465	1450 6477	-0.3011		73	0.0012	÷.	P. NEVSDADUDASAD. 0				
	-	133	499.1349	1494.3828	1494.6694	-0.2866	0	92	0.0011	î	L.DPSLMEMTEAALR.L + 2 Oxidation (M)				
	~	145	526.1538	1575,4396	1575,7814	-0.3418	0	(61)		1	R.ALTETIMEDDAIER.A + [-48.0000 at F8]				
	~	156	820.7283	1639.4420	1639.7763	-0.3343	0	97	5.1e-06	1	R.ALTETINFDDAIER.A + Oxidation (M)				
i i	~	165	841.2310	1680.4474	1680.8029	-0.3554	0	(75)		1	R_ALTETINFDDAIER.A + Oxidation (M); [+41.0266 at N-term A]				
6	~	170	864.2888	1726.5629	1726.9294	-0.3664	0	44	0.9	1	K.AYTVLLYGNGPGYVLK.D				
6	~	176	879.2425	1756.4705	1756.8420	-0.3715	0	83		1	G.IIPVEEENPDFWNR.E				
		204	956.2437	1910.4729	1910.8601	-0.3872	0	29	28	3	R.DSTLDPSLMEMTEAALR.L + 2 Oxidation (M)				
6	~	208	975.8100	1949.6055	1950.0245	-0.4190	0	85	6.6e-05	1	K.NLIIFLGDGMGVSTVTAAR.I + Oxidation (M)				
6	~	209	976.2340	1950.4534	1950.8555	-0.4021	0	(27)	42	1	K.DGARPDVTESESGSPEYR.Q				
6	~	211	656.1752	1965.5039	1964.8712	0.6327	0	(72)		1	K.DGARPDVTESESGSPEYR.Q + [+14.0157 at T8]				
	_	213	664.5518	1990.6336	1991.0510	-0.4174	0	(58)		4	K_NLIIFLGDGMGVSTVTAAR.I + Oxidation (M); [+41.0266 at N-term N]				
	~	216	1001.2027	2000.3908	2000.8058	-0.4150	0	(65)	0.0069	1	R.MGTPDPEYPDDYSQGGTR.L + Oxidation (M)				
	× 1	217	667.8046	2000.3919	2000.8058	-0.4139		70	0.002	1	K.MUTPPETPUTSUUGR.L + UXIdation (M)				
	× 1	222	691 9205	2007.4466	2007.8770	-0.4304		75		1	N_DUMARDVIESESUSTEIK.V + [+37.V213 at N-TeTM J] P.MCTPDPEVPDDVSCCCTP I + Acetyl (N-term): Ovidation (N): [-0.9840 at F2]				
	-	252	784.5440	2350.6103	2351.1030	-0.4927		(61)		1	REMOVE DEETHAGE DVAVEAR.G + $[-17,0265]$ at N-term 01				
	-	253	790.2187	2367.6341	2368.1295	-0.4954		94	7.6e-06	÷.	R. OOSAVPLDEETHAGEDVAVFAR. G				
	~	2 60	809.2208	2424.6406	2425.1510	-0.5104	0	(66)		1	R.00SAVPLDEETHAGEDVAVFAR.G + [+57.0215 at N-term 0]				
	~	275	920.5878	2758.7415	2759.3582	-0.6167	0	90		1	R.QEGCQDIATQLISNMDIDVILGGGR.K + Acetyl (N-term); Oxidation (M); [-9.9476 at				
		_													
<											>				
۲											🗟 🔮 Internet				
	A .	۸ C	COT	• • • •	adifia	ation	•				MATRIX				
		-12		: ///	ounc	auon	5				© 2007-2012 Matrix Science				
L															

And here is the first hit of the results report. The additional matches, found in the error tolerant search, are the ones without expect values. One of these, query 133, is a simple, non-specific peptide with a very good score. There's another example for query 176. The error tolerant search is a much better way of picking up non-specific peptides than searching the entire database with semi-trypsin or no enzyme. We only fail to get such matches in an error tolerant search if there are no matches to the protein in the first pass search. However, you have to ask yourself whether you would believe a protein hit in which the only peptide match was non-specific. I think the answer is no.



The matches from an error tolerant search are aggressively filtered to remove junk matches

🗿 Pe	ptide	Summ	ary Report (Er	rror tolerant e	kample) - Micro	osoft Intern	et Exp	lorer							
File	Edit	View	Favorites To	ools Help							a da fa se a companya da 🦓				
G	Back	• €) · 🖹 💈	🖌 🏠 🖉 s	earch 🤺 Favo	rites 🧭	@ •	2	🔁 •	Pov	vermarks 膬 🔺 ۶				
Addre	ss 🍓	http://	www.matrixscienc	ce.com/cgi/master	results.pl?file=/e	data/20070626	/FoGcri	e5.dat			💌 🔁 Go				
											,				
Se	elect.	All	Select None	Searc	h Selected	Еггог	tolera	nt							
· .		11517	Mage	56371	Score: 782	Queries	mat	ched.	27 emPl	ατ • ο	78				
	P	UMALP	PA NID: - 1	Homo sapien:	8	•									
		heck	to include	this hit in	n error tole	erant sear	ch								
	Query Observed Mr(expt) Mr(calc) Delta Miss Score Expect Rank Peptide														
	_0	uery	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Expect	Rank	Peptide				
	⊻	27	462.6807	923.3468	923.5116	-0.1649		33	16	1	R.FPIVALSK.T				
		52	564 6804	1032.3373	1127.5764	-0.2229	0	10	2.8e+03	5	R. GSSIFGLAFGK.A				
		65	567.6567	1133.2987	1133.5499	-0.2511	0	44	1.1	1	$R_{\rm GNEVISVOR,A} + 0xidation (M)$				
		86	614.2001	1226.3856	1226.6329	-0.2473	0	28	41	2	K.LGPEIPLAMDR.F + Oxidation (M)				
	V	100	653.2101	1304.4057	1304.6837	-0.2780	0	(87)	5.7e-05	1	K. GNFQTIGLSAAAR.F				
	v	124	710.2235	1418.4324	1418.7154	-0.2829	0	95		1	K. <u>G</u> NFQTIGLSAAAR.F + Acetyl (N-term); [+72.0211 at N-term G]				
	V	126	726.1806	1450.3465	1450.6477	-0.3011	0	73	0.0012	1	R. NWYSDADVPASAR. Q				
	V	133	499.1349	1494.3828	1494.6694	-0.2866	0	92		1	L.DPSLMEMTEAALR.L + 2 Oxidation (M)				
	v	145	526.1538	1575.4396	1575.7814	-0.3418	0	(61)		1	R.ALTETINFDDAIER.A + [-48.0000 at F8]				
	v	156	820.7283	1639.4420	1639.7763	-0.3343	0	97	5.1e-06	1	R.ALTETIMFDDAIER.A + Oxidation (M)				
	✓	165	841.2310	1680.4474	1680.8029	-0.3554	0	(75)		1	R_ALTETIMEDDAIER.A + Oxidation (M); [+41.0266 at N-term A]				
	✓	170	864.2888	1726.5629	1726.9294	-0.3664	0	44	0.9	1	K.AYTVLLYGNGPGYVLK.D				
	✓	176	879.2425	1756.4705	1756.8420	-0.3715	0	83		1	G. IIPVEEENPDFWNR.E				
		204	956.2437	1910.4729	1910.8601	-0.3872	0	29	28	3	K.DSTLDPSLARMTERALR.L + 2 Oxidation (M)				
		200	975.8100	1949.6033	1950.0245	-0.4190		(27)	6.68-03	1	K.RLIIILGDOMGVSIVIAAK.I + OXIGATION (M)				
		211	656.1752	1965.5039	1964.8712	0.6327		(72)		- 1	K.DCARPDVTESESGSPEVR.0 + [+14.0157 at T8]				
		213	664.5518	1990.6336	1991.0510	-0.4174	0	(58)		4	K.NLIIFLGDGMGVSTVTAAR.I + Oxidation (M); [+41.0266 at N-term N]				
	v	216	1001.2027	2000.3908	2000.8058	-0.4150	0	(65)	0.0069	1	R.MGTPDPEYPDDYSQGGTR.L + Oxidation (M)				
	v	217	667.8046	2000.3919	2000.8058	-0.4139	0	70	0.002	1	R.MGTPDPEYPDDYSQGGTR.L + Oxidation (M)				
	v	218	670.1561	2007.4466	2007.8770	-0.4304	0	75		1	K_DGARPDVTESESGSPEYR.Q + [+57,0215 at N-term D]				
	V	222	681.8205	2042.4397	2041.8324	0.6073	0	(61)		1	R_MGTPDPEYPDDYSQGGTR.L + Acet (N-term); Oxidation (M); [-0.9840 at E7]				
	v	252	784.5440	2350.6103	2351.1030	-0.4927	0	(69)		1	R_QQSAVPLDEETHAGEDVAVFAR.G + [Possible Assignments:				
	v	253	790.2187	2367.6341	2368.1295	-0.4954	0	94	7.6e-06	1	R. QQSAVPLDEETHAGEDVAVFAR. G				
	V	260	809.2208	2424.6406	2425.1510	-0.5104	0	(66)		1	R_QQSAVPLDEETHAGEDVAVFAR.G + [Carbamidomethyl (N-term) [+57.0215]				
	<	275	920.5878	2758.7415	2759.3582	-0.6167	0	90		1	R_QEGCQDIATQLISNMDIDVILGGGR.K Carboxymethyl (N-term) [+58.0055] .9476 at :				
_															
<u>s</u>															
2 1:A	AA517	08 2:AA	AH09647 3:CAJ15	5103 4:512076 5:	AAA51709 8:AAA	98616		_		_	S Internet				
N	A	45	SCO ⁻	F : <i>M</i>	odific	ation	s				© 2007-2012 Matrix Science				

Take a look at the match to query 218. The mass tolerance for this search was fairly wide, so the observed mass difference could correspond to either carbamidomethylation or carboxymethylation at the N-terminus. Since this sample was alkylated with iodoacetamide, we would choose carbamidomethylation as the more likely suspect, especially as this brings the error on the precursor mass into line with the general trend, whereas carboxymethylation would give an error of +0.6 Da. The assignment to carbamidomethylation is also very believable, because this is a known artefact of over-alkylation. The same modification is found for query 260.

🗿 Pept	ide Su	mma <mark>ry Report (</mark> E	rror tolerant e	xample) - Micro	osoft Intern	et Exp	lorer								
File	Edit \	view Favorites To	ools Help							Na katala 🔊					
()	Back 🔹	۲ 🖹 🕄) 🏠 🔎 s	iearch 🤺 Favo	rites 🧭	@ •	2	🔁 -	Poi	vermarks 🌃 🔺 %					
Address	🕘 ht	tp://www.matrixscien	ce.com/cgi/master	_results.pl?file=/e	data/20070626	/FoGcrie	eS.dat			💌 🄁 Go					
										<u>^</u>					
Sele	ect All	Select None	e Searc	ch Selected	Еггог	tolera	nt								
l		51709 Mage	. 56221	Score: 702	Querier		ched	27 emP/		79					
1 .	HUM	ALPPA NID: - 1	Homo sapien	30010. 702	Querres	s mac	cheu.	27 CHEZ	u. 0	. 10					
6	Check to include this hit in error tolerant search														
	Duary Obsawad Ny(amt) Ny(ala) Balta Nice Cova Pomet Dank Dantida														
	Quer	y Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Expect	Rank	Peptide					
E	2 2	462.6807	923.3468	923.5116	-0.1649	0	33	16	1	R.FPYVALSK.T					
	3	1 517.1760	1032.3375	1032.5604	-0.2229	U	70	0.0036	3	R. GSSIFGLAPGK.A					
	a -	567 6567	1127.3403	1127.5709	-0.2501	0	10	2.00+03	1	P CHEVISION A + Ovidation (M)					
	2 <u>2</u>	6 614.2001	1226.3856	1226.6329	-0.2473	0	28	41	2	K.LGPEIPLANDR.F + Oxidation (M)					
L R	10	653.2101	1304.4057	1304.6837	-0.2780	0	(87)	5.7e-05	1	K. GNFOTIGLSAAAR.F					
6	12	4 710.2235	1418.4324	1418.7154	-0.2829	0	95		1	K.GNFQTIGLSAAAR.F + Acetyl (N-term); [+72.0211 at N-term 6]					
	12	6 726.1806	1450.3465	1450.6477	-0.3011	0	73	0.0012	1	R. NWYSDADVPASAR. Q					
	13	499.1349	1494.3828	1494.6694	-0.2866	0	92		1	L.DPSLMEMTEAALR.L + 2 Oxidation (M)					
E	14	526.1538	1575.4396	1575.7814	-0.3418	0	(61)		1	R.ALTETIMEDDAIER.A + [-48.0000 at F8]					
E	15	6 820.7283	1639.4420	1639.7763	-0.3343	0	97	5.1e-06	1	R.ALTETIMEDDAIER.A + Oxidation (M)					
5	 16 	841.2310	1680.4474	1680.8029	-0.3554	0	(75)		1	R_ALTETIMEDDAIER.A + Oxidation (M); [+41.0266 at N-term A]					
6	17	864.2888	1726.5629	1726.9294	-0.3664	0	44	0.9	1	K.AYTVLLYGNGPGYVLK.D					
E	17	879.2425	1756.4705	1756.8420	-0.3715	0	83		1	G.IIPVEEENPDFWNR.E					
	20	956.2437	1910.4729	1910.8601	-0.3872	0	29	28	3	R.DSTLDPSLMEMTEAALR.L + 2 Oxidation (M)					
	20	975.8100	1949.6055	1950.0245	-0.4190	0	85	6.6e-05	1	K.NLIIFLGDGMGVSTVTAAR.I + Oxidation (M)					
	20	976.2340	1950.4534	1950.8555	-0.4021	0	(27)	42	1	K.DGARPDVTESESGSPEYR.Q					
6	21	<u>1</u> 656.1752	1965.5039	1964.8712	-0 4174	U	(72)		1	K.DURKPDVIESESUSPEIR.U + $[+14.0157]$ at T8] K.W.TIFLEDCHERVSTUTARD I + Oridation (M): [+41.0266 at N-term N]					
	21	<u>6</u> 1001.3310	2000.3908	2000.8058	-0.4150	0	(65)	0.0069	1	R.MCTPDPEVPDDVSOGGTR.L. + Oxidation (N)					
	21	7 667.8046	2000.3919	2000.8058	-0.4139	0	70	0.002	-	R.MGTPDPEYPDDYSOGGTR.L + Oxidation (M)					
	21	8 670.1561	2007.4466	2007.8770	-0.4304	0	75		1	K.DGARPDVTESESGSPEYR.Q + [+57.0215 at N-term D]					
	22	2 681.8205	2042.4397	2041.8324	0.6073	0	(61)		1	R.MGTPDPEYPDDYSQGGTR.L + Acetyl (N-term); Oxidation (M); [-0.9840 at E7]					
6	25	2 784.5440	2350.6103	2351.1030	-0.4927	0	(69)		1	R_QQSAVPLDEETHAGEDVAVFAR.G + [-17.0265 at N-term Q]					
E	25	3 790.2187	2367.6341	2368.1295	-0.4954	0	94	7.6e-06	1	R. QQSAVPLDEETHAGEDVAVFAR. G					
E	26	809.2208	2424.6406	2425.1510	-0.5104	0	(66)		1	R_QQSAVPLDEETHAGEDVAVFAR.G + [+57.0 Possible Assignments:					
E	27	920.5878	2758.7415	2759.3582	-0.6167	0	90		1	R.QEGCQDIATQLISNMDIDVILGGGR.K + Act					
										Gin->pyro-Giu (N-Cerm Q) [-17.0285]					
<										8					
🕘 1:AAA	451708	2:AAH09647 3:CA315	5103 4:512076 5:	AAA51709 8:AAA	98616					🛜 🔮 Internet					
N	٨A	SCO	T : <i>M</i>	odific	ation	s				© 2007-2012 Matrix Science					

Another easily believable assignment is pyro-Glu for the match to query 252.

🕲 Peptide Summary Report (Error tolerant example) - Microsoft Internet Explorer 📃 💽 🔀											
File Edk View Favorites Tools Help											
A part - A - R & A Cauch Stranger A A. B. Comments 24 4 24											
Constructs of the construction of the construc											
Address	, Address 🍘 http://www.matrixsclence.com/cgi/matter_results.pl/file=/data/20070626/jFoGcrie5.dat 🔤 🔂 600										
Select	Select All Select None Search Selected Error tolerant										
1.	1. AAAS1708 Hass: 56371 Score: 782 Queries matched: 27 emPAI: 0.78 HUNLIPAIND: - Homo saminan										
	nonAurra NID: - nomo septens										
	uery	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Expect	Rank	Peptide	
v	27	462.6807	923.3468	923.5116	-0.1649	0	33	16	1	R.FPYVALSK.T	
	41	517.1760	1032.3375	1032.5604	-0.2229	0	70	0.0036	3	R.GSSIFGLAPGK.A	
_	<u>62</u>	564.6804	1127.3463	1127.5764	-0.2301	0	10	2.8e+03	6	R.GFFLFVEGGR.I	
	<u>65</u>	567.6567	1133.2987	1133.5499	-0.2511	0	44	1.1	1	R.GREVISVHIR.A + Oxidation (M)	
	100	653 2101	1226.3836	1226.6329	-0.2473		(87)	5 70-05	1	K CHEATLEI SANAD F	
	124	710.2235	1418.4324	1418.7154	-0.2829		95	5.76-05	÷.	K.GNFOTIGLSAMAR.F + Acetyl (N-term): $[+72,0211]$ at N-term Gl	
	12.6	726.1806	1450.3465	1450.6477	-0.3011	0	73	0.0012	- 1	R. NWYSDADVPASAR. 0	
	133	499.1349	1494.3828	1494.6694	-0.2866	0	92		1	L.DPSLMEMTEAALR.L + 2 Oxidation (M)	
	145	526.1538	1575.4396	1575.7814	-0.3418	0	(61)		1	R.ALTETINEDDAIER.A + [-48.0000 at F8]	
v	156	820.7283	1639.4420	1639.7763	-0.3343	0	97	5.1e-06	1	R.ALTETIMFDDAIER.A + Oxidation (M)	
V	165	841.2310	1680.4474	1680.8029	-0.3554	0	(75)		1	R_ALTETIMFDDAIER.A + Oxidation (M); [+41.0266 at N-term A]	
V	170	864.2888	1726.5629	1726.9294	-0.3664	0	44	0.9	1	K.AYTVLLYGNGPGYVLK.D	
V	176	879.2425	1756.4705	1756.8420	-0.3715	0	83		1	G.IIPVEEENPDFWNR.E	
_	<u>204</u>	956.2437	1910.4729	1910.8601	-0.3872	0	29	28	3	R.DSTLDPSLMEMTEAALR.L + 2 Oxidation (M)	
	208	975.8100	1949.6055	1950.0245	-0.4190	0	85	6.6e-05	1	K.NLIIFLGDGMGVSTVTAAR.I + Oxidation (M)	
	209	976.2340	1950.4534	1950.8555	-0.4021	0	(27)	42	1	K.DGARPDVTESESGSPEYR.Q	
	211	656.1752	1965.5039	1964.8712	0.6327	0	(72)		1	K.DUARPDVIESESUSPEIR.U + $[+14.0157]$ at 18] K.W.TIELEDCHCHCSTUTAAR I + Oxidat ^(h) (M) + [+41.0266] at N-term NI	
	216	1001.2027	2000.3908	2000.8058	-0.4150	0	(65)	0.0069	1	R.MCTPDPEVPDDVSOGGTR L + Ovidatio	
	217	667.8046	2000.3919	2000.8058	-0.4139	0	70	0.002	1	R.MGTPDPEYPDDYSQGGTR.L + Oxidatic	
	218	670.1561	2007.4466	2007.8770	-0.4304	0	75		1	K.DGARPDVTESESGSPEYR.Q + [+57.02] Thr->&sn (T) [+12.9952]	
	222	681.8205	2042.4397	2041.8324	0.6073	0	(61)		1	R.MGTPDPEYPDDYSQGGTR.L + Acetyl Methylamine (T) [+13.0316] .9840 at E7]	
V	252	784.5440	2350.6103	2351.1030	-0.4927	0	(69)		1	R_QQSAVPLDEETHAGEDVAVFAR.G + [-1]	
v	253	790.2187	2367.6341	2368.1295	-0.4954	0	94	7.6e-06	1	R.QQSAVPLDEETHAGEDVAVFAR.G	
V	260	809.2208	2424.6406	2425.1510	-0.5104	0	(66)		1	R_QQSAVPLDEETHAGEDVAVFAR.G + [+57.0215 at N-term Q]	
V	<u>275</u>	920.5878	2758.7415	2759.3582	-0.6167	0	90		1	R_QEGCQDIATQLISNMDIDVILGGGR.K + Acetyl (N-term); Oxidation (M); [-0.9476 t]	
🛃 IRAADI./UD 2/ANTI/DRV/ 37LAIIDIUS 1/SIL/UD SIAAAS1709 8/AAAS8616											
	MATRIX										
M	A:	SUU	: M	odific	ation	S				© 2007-2012 Matrix Science	
Join Science											

As is methylation ay T8 for query 211

🔋 Peptide Summary Report (Error tolerant example) - Microsoft Internet Explorer 🔹 💽 🔀											
File Edit View Favorites Tools Help											
😮 Back + 💿 - 🖹 🖉 🏠 🔎 Search 🔆 Favorites 🤣 🔗 - 😓 🛱 * Povernaris 🥻 🔺 🛠											
Address	Address 👔 http://www.matrixscience.com/cgi/master_results.pl/file=jdata/2007/0626/FoGories.dat 👻 🔁 Go										
Selec	Select All Select None Search Selected Trior tolerant										
1.	1. AAA51708 Mass: 56371 Score: 782 Oueries matched: 27 emPAI: 0.78										
HUMLEPA NDI - Homo sapiens											
Check to include this hit in error tolerant search											
	Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Expect	Rank	Peptide	
	41	462.6007	923.3466	923.5116	-0.2229	0	33	0.0036	3	R. FFIVALSA. I D. GSSTEGLADEK A	
	62	564.6804	1127.3463	1127.5764	-0.2301	0	10	2.8e+03	6	R. GFFLFVEGGR. I	
	65	567.6567	1133.2987	1133.5499	-0.2511	0	44	1.1	1	R.GNEVISVMNR.A + Oxidation (M)	
_	86	614.2001	1226.3856	1226.6329	-0.2473	0	28	41	2	K.LGPEIPLAMDR.F + Oxidation (M)	
	100	653.2101	1304.4057	1304.6837	-0.2780	0	(87)	5.7e-05	1	K. GNFQTIGLSAAAR.F	
	124	710.2235	1418.4324	1418.7154	-0.2829	0	95		1	K. <u>GNFQTIGLSAAAR.F</u> + Acetyl (N-term); [+72.0211 at N-term 6]	
	126	726.1806	1450.3465	1450.6477	-0.3011	0	73	0.0012	1	R. NWYSDADVPASAR. Q	
	<u>133</u>	499.1349	1494.3828	1494.6694	-0.2866	0	92		1	L.DPSLMEMTEAALR.L + 2 Oxidation (M)	
	145	526.1538	1575.4396	1575.7814	-0.3418	0	(61)		1	R.ALTETIMEDDAIER.A + [-48.0000 at F8]	
	156	820.7283	1639.4420	1639.7763	-0.3343	0	97	5.1e-06	1	R.ALTETINFDDAIER.A + Oxidati_h (N)	
	<u>165</u>	841.2310	1680.4474	1680.8029	-0.3554	0	(75)		1	R_ALTETIMFDDAIER.A + Oxidati Possible Assignments: m A]	
	170	864.2888	1726.3629	1726.9294	-0.3554		44	0.9	1	C. TIMERENDERAD F. Phe->Val (F) [-48.0000]	
	204	879.2423	1910 4729	1910 8601	-0.3/15	0	20	28	3	D DETUDES MENTERALD I + 2 0	
	208	975.8100	1949.6055	1950.0245	-0.4190	0	85	6.66-05	1	K.NLITELEDEMENSTUTAAR.I + Oxidation (M)	
	209	976.2340	1950.4534	1950.8555	-0.4021	0	(27)	42	1	K.DGARPDVTESESGSPEYR.0	
	211	656.1752	1965.5039	1964.8712	0.6327	0	(72)		1	K.DGARPDVTESESGSPEYR.Q + [+14.0157 at T8]	
	213	664.5518	1990.6336	1991.0510	-0.4174	0	(58)		4	K_NLIIFLGDGMGVSTVTAAR.I + Oxidation (M); [+41.0266 at N-term N]	
	216	1001.2027	2000.3908	2000.8058	-0.4150	0	(65)	0.0069	1	R.MGTPDPEYPDDYSQGGTR.L + Oxidation (M)	
	217	667.8046	2000.3919	2000.8058	-0.4139	0	70	0.002	1	R.MGTPDPEYPDDYSQGGTR.L + Oxidation (M)	
	218	670.1561	2007.4466	2007.8770	-0.4304	0	75		1	K_DGARPDVTESESGSPEYR.Q + [+57.0215 at N-term D]	
	222	681.8205	2042.4397	2041.8324	0.6073	0	(61)		1	R.MGTPDPEYPDDYSQGGTR.L + Acetyl (N-term); Oxidation (M); [-0.9840 at E7]	
	252	784.5440	2350.6103	2351.1030	-0.4927	0	(69)		1	R_QQSAVPLDEETHAGEDVAVFAR.G + [-17.0265 at N-term Q]	
	253	790.2187	2367.6341	2368.1295	-0.4954		94	7.66-06	1	R. QUSAVPLDEETHAGEDVAVPAR. G	
	200	920.5878	2424.0400	2425.1510	-0.5104	0	(00)		1	$R_{\rm v}$ (VSAVPLDEE INAGEDVAVPAR. 6 + (+37.0215) at N-term ()	
	,	52010010		210510002	010101				-	A A A A A A A A A A A A A A A A A A A	
<	<x< td=""></x<>										
2) 1.4AA51708 2.4A4509647 3:CA15103 4:512076 5:AAA51709 8:AAA68616											
N	MASCOT : Modifications © 2007-2012 Matrix Science										

In other cases, the match may be good, but the assignment is not believable. Query 145 is listed with a substitution at F8 causing a loss of 48 Da. This seems unlikely because we have 2 other matches to the same peptide without any substitution. What else could it be? Well, notice that the other two matches are both oxidised at M7. If we suppose this peptide is also oxidised, then the mass shift becomes -64, which is a well-known loss for oxidised methionine, (loss of methanesulfenic acid). This would seem a much more likely explanation for this match.

It is important to understand that the error tolerant search finds new matches by introducing mass shifts at different positions in the database sequences. The match may be very strong, but figuring out a credible assignment can require a bit of detective work.

Peptide Summary Report (Error tolerant example) - Microsoft Internet Explorer										
File Edit Wew Favorites Tools Help										
🕝 Back · 🕘 · 🙀 🖉 🏠 🔎 Search 🐈 Favorites 🤣 🔗 - 🌺 🧖 · Powermarks 🐲 🚸										
Address 🕘 http://www.matrixscience.com/cgi/master_results.pl?file=/data/20070626/FoGcrie5.dat	🛩 🛃 Go									
11 INTO Mage: 22070 Score: 454 Querieg metched: 16 amDAT: 1.42										
trypsin (EC 3.4.21.4) (isopropylphosphorylated) - bovine										
Check to include this hit in error tolerant search										
Query Observed Mr(expt) Mr(calc) Delta Miss Score Expect Rank	Peptide									
✓ <u>71</u> 577.1685 1152.3225 1152.5663 -0.2438 0 87 4.6e-05 1	K.SSGTSYPDVLK.C									
	K.VCNYVSWIK.Q									
✓ <u>78</u> 598.1756 1194.3366 1194.5768 -0.2402 0 (69) 1	K. <u>SSGTSYPDVLK.C + [+42.0106</u> at N-term S]									
83 606.1832 1210.3559 1210.5717 -0.2158 0 (61) 1	$K_{SSGTSYPPVLK,C} + [+38.0055] \text{ at } N-\text{term S}]$									
V <u>54</u> 040.1270 1270.2411 1270.4023 -0.2219 0 (07) 1 132 745.7224 1489.4302 1489.7348 -0.3046 0 72 0.0017 1	K.I.OGTUSWESGCAOK N									
▼ 229 1081.7685 2161.5224 2162.0491 -0.5267 0 156 5.1e-12 1	R.LGEDNINVVEGNEOFISASK.S									
✓ 230 721.5398 2161.5976 2162.0491 -0.4515 0 (94) 8.2e-06 1	R.LGEDNINVVEGNEQFISASK.S									
V 231 721.8998 2162.6775 2162.0491 0.6284 0 (42) 1.5 1	R.LGEDNINVVEGNEQFISASK.S									
	R.LGEDNINVVEGNEQFISASK.S + [+23.9748 at N-term L]									
✓ 234 1094.8114 2187.6082 2188.0284 -0.4201 0 (97) 1	R.LGEDNINVVEGNEQFISASK.S + Acetyl (N-term); [-16.0313 at N-term L]									
✓ 236 1102.8029 2203.5912 2204.0961 -0.5048 0 (102) 1	R.LGEDNINVVEGNEQFISASK.S + [+42.0470 at G2]									
735.3400 2203.3903 2204.0397 -0.4614 0 (67) 0.0043 1 7 Ton scoring mentide matches to query 236	R_LOEDRINVVEGREGFISASK.S + ACCCY1 (R-CCIM) B I GEDNINRVEGREGFISASK S + [+57, 0215 at C-term K]									
2 142: Sum of 8 scans in range 2405 (rt=2102.48) to 2426 (rt=2116.32)	R.LGEDNINVVEGNEOFISASK.S + [+57.0215 at C-term K]									
	K.SIVHPSYDSNTLNNDDMLIK.L									
102.4 -0.5048 11+ 1NTP R.LGEDNINVVEGNEOFISASK.S										
12 TP 100.8 1.7e-06 -0.4684 11+ 1NTP R_LGEDNINVVEGNEQFISASK.S	7									
12. 17.0 4e+02 -0.5677 R_VGDPFNPKVTVGPVNNPGQVK.Y										
Che 13.0 1e+03 0.5940 K.GGARVGWIVVCHGEGMMEDK.S										
12.9 1e+03 -0.4395 K_VLSCDYVDQSSNLTIFSSK.E										
Que 12.8 1.12+03 0.3588 R.DPNSPKVSAVSAIVNKGLPLK.A 12.3 1.2e+03 -0.5412 K.TPTGPNASSSAVPSSKYTVAIK.D	Peptide									
✓ 11.9 1.3e+03 0.4783 M.RVLTLNDKDLFMAHDVMK.T	K. SIVHPSYNSN. T									
11.8 1.4e+03 0.2933 R_YPQLPIVGLVPALKPAISASK.T 11.5 1.4e+03 -0.4837 R.OAFVKPEDIDVIXAHGSGTK.O	K.SSGTSYPDVLK.C									
	K.SSGTSYPDVLK.C + [+42.0106 at N-term S]									
<u>83</u> 606.1852 1210.3559 1210.5717 -0.2158 0 (61) 1	K_SSGTSYPDVLK.C + [+58.0055 at N-term S]									
<u>94</u> 640.1278 1278.2411 1278.4629 -0.2219 0 (67) 1	K.SSGTSYPDVLK.C + [+125.8966 at Y6]									
<u>132</u> 745.7224 1489.4302 1489.7348 -0.3046 0 72 0.0017 1	K.LQGIVSWGSGCAQK.N									
229 1081.7685 2161.5224 2162.0491 -0.5267 0 156 5.1e-12 1	K. LOEDNINVVEGNEUP ISASK. S									
11:INTP 12:TRBOTR	S Diternet									
MASCOT : Modifications	© 2007-2012 Matrix Science									
	SCIENCE									

You should also look at the other yellow pop-up when trying to decide whether to accept a match or not. In this example, the error tolerant search was able to get a slightly higher score by shifting a modification of +42 Da from the amino terminus to the adjacent glycine. However, as score increase of 2 in 100 is negligible. Much more believeable to take the original match from the first pass search, which can be explained as N-terminal acetylation.



In summary, an error tolerant search

•Can successfully locate mass differences corresponding to a single unsuspected modification or a single SNP per peptide

•User must decide on best explanation for the observed differences

•Limited to proteins which have at least one good peptide match ... not very useful for (say) MHC peptides