Lawrence Berkeley National Laboratory
Recent Work

Title
DNA Based Pair Substitutions Detected with MALDI-TOF Mass Spectrometry

Permalink
https://escholarship.org/uc/item/1jn1t1g6

Authors
Benner, W.H.
Jaklevic, J.M.

Publication Date
1995-03-01

eScholarship.org Powered by the California Digital Library
University of California

LBL-36870
UC-406
Preprint

Lawrence Berkeley Laboratory

UNIVERSITY OF CALIFORNIA

Engineering Division

Submitted to Nucleic Acids Research

DNA Base-pair Substitutions Detected with MALDI-TOF Mass Spectrometry

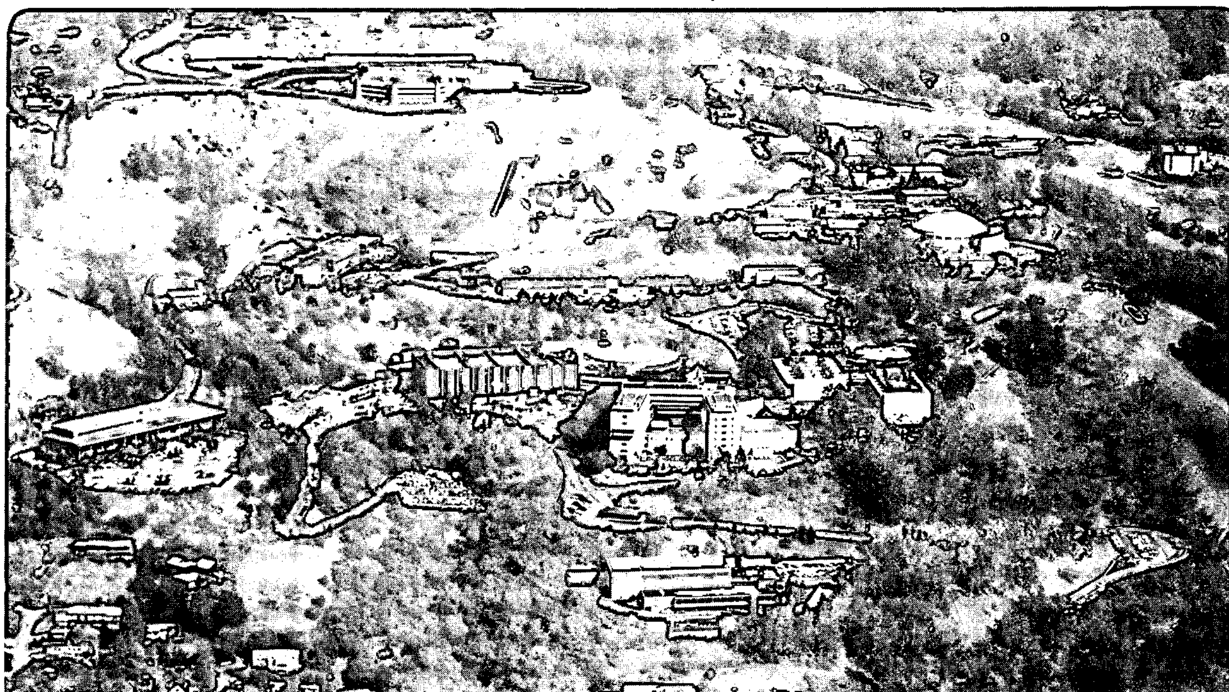
W.H. Benner and J.M. Jaklevic

March 1995

U. C. Lawrence Berkeley Laboratory
Library, Berkeley

FOR REFERENCE

Not to be taken from this room



| REFERENCE COPY |
| Does Not | Copy 1
| Circulate |
. Bldg. 50 Library.
LBL-36870

Prepared for the U.S. Department of Energy under Contract Number DE-AC03-76SF00098

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

LBL-36870
UC-406

"DNA Base-pair Substitutions Detected with MALDI-TOF Mass Spectrometry"

W.H. Benner, J.M. Jaklevic

Human Genome Center Instrumentation Group
Lawrence Berkeley Laboratory
University of California
Berkeley, CA 94720

March 1995

This work was supported by the Director, Office of Energy Research, Office of Biological and Environmental Research, Analytical Technology Division, of the U.S. Department of Energy under Contract DE-AC03-76SF00098.

DNA Base-pair Substitutions Detected in dsDNA with MALDI-TOF Mass Spectrometry

W. H. Benner and J. Jaklevic
Human Genome Center Instrumentation Group
Engineering Research and Development Department
Engineering Division
Lawrence Berkeley Laboratory,
University of California, Berkeley, CA 94720

Abstract

We demonstrate the utility of matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS) for identifying the presence or absence of single base-pair substitutions in a short section of the cystic fibrosis gene. The analysis was performed in two ways: (1) the molecular weight of known lengths of ssDNA from a section of this gene was determined by mass spectrometry, and (2) short complementary strands from the identical section of the gene were denatured and the difference in molecular weight of the strands was determined. Oligonucleotides corresponding to the wild type sequence were synthesized. The G551D or the R553X mutations were created by introducing single-base substitutions in this oligonucleotide. We recreated dsDNA by annealing each of these three oligonucleotides to complementary oligonucleotides. The MALDI mass spectrometric technique, operating at a mass resolution approaching 1000, correctly identified five of these six oligonucleotides. The most reliable identification was obtained by measuring the difference in molecular weight of the denatured strands. The difference in molecular weight of the complementary strands comprising the wild type, the G551D, and the R553X is 52.2 Da, 83.2 Da, and 21.2 Da, respectively. The MALDI analysis shows promise as a way to rapidly detect the presence of base-pair substitutions without the need for sequencing.

Key words: MALDI, mass spectrometry, time-of-flight, cystic fibrosis gene, DNA

Introduction

DNA diagnostics comprises an assortment of techniques to detect not only genetic diseases[1-2], but also the presence of pathogenic organisms such as bacteria[3] and viruses[4] in body fluids. The technology relies on the discrimination of DNA sequences known to be unique to individuals or groups of organisms. Chromosomal markers, restriction fragment length polymorphisms and DNA sequence analysis illustrate a few of the ways that disease genes and DNA isolated from pathogens might be recognized. Of these techniques, sequencing reliably identifies mutations or sequences specific to an organism. Genetic disease diagnosis by means of sequencing is performed by selecting a region of a gene where the sequence is known to encompass previously identified mutations. Mutations might occur at random locations along a gene, but the repeated occurrence of a base deletion(s) or substitution(s) at the same location in a gene has correlated with particular forms of genetic disease. The G551D and R553X substitutions in the cystic fibrosis gene contribute two forms of cystic fibrosis in humans. A variety of biochemical techniques can be used to isolate and amplify such regions of interest, and gel electrophoresis is subsequently used to separate amplified DNA into bands that identify the presence or absence of the region of interest or the specific sequence. Gel electrophoresis is customarily used to separate and identify sequencing reaction products; it works remarkably well but the procedure is time consuming. Mass spectrometric analyses might provide an alternative to gel electrophoresis because it provides results quickly.

Mass spectrometric techniques continue to be attractive for determining the molecular weight of large molecules, such as proteins and DNA fragments. Current difficulties associated with the generation of unfragmented molecular ions of DNA limit the applicability of MALDI-TOF-MS technology to the determination of the molecular weight of relatively small DNA molecules.[5] MALDI-TOF-MS sequencing procedures also have been reported by Pieles *et al.*[6] for short oligonucleotides. Recently, we developed a mass spectrometer technique that showed the application of MALDI-TOF-MS for determining the difference in molecular weight of the complementary strands of denatured DNA.[7] Using this procedure, we report now the application of MALDI-TOF-MS to detect base-pair substitutions in short strands of dsDNA without the need for sequencing reactions. The procedure relies on the determination of the molecular weight of each of the complementary strands of dsDNA following heat denaturation. We demonstrate the application of this approach to the identification of two mutations in the cystic fibrosis gene.

The G551D and the R553X mutations in the cystic fibrosis gene occur in a 32 base-long region which has the following sequence:

5'- cac TGA GTG GAG GTC AAC GAG CAA Gaa tt

3'- act CAC CTC CAG TTG CTC GTT CTT AAa ga -5'

in which the capital letters designate wild type (WT). As written, the top strand is WT+ and the lower is WT-. Two different endonucleases cut at sequences corresponding to the 5' overhangs on each end of the dsDNA, thus providing a way to isolate the region of interest from genomic DNA. The underlined letters in the sequences below indicate bases substituted at the same position in the WT allele. For example, adenosine substitutes for guanosine at the tenth base from the 5' end of WT+ in the 551+ oligonucleotide. The 551+/- complementary strands correspond to the G551D mutation and 553+/- identifies the R553X mutation.

WT+:	5'-TGA GTG GAG GTC AAC GAG CAA G-3'	22 bases, 6873.4 Da
551+:	5'-TGA GTG GAG ATC AAC GAG CAA G-3'	22 bases, 6857.4 Da
WT-:	5'-AAT TCT TGC TCG TTG ACC TCC AC-3'	23 bases, 6925.6 Da
551-:	5'-AAT TCT TGC TCG TTG ATC TCC AC-3'	23 bases, 6940.6 Da
553-:	5'-AAT TCT TGC TCA TTG ACC TCC AC-3'	23 bases, 6909.6 Da

Using two analytical approaches, we attempted to confirm that MALDI-TOF-MS could determine the occurrence of the base substitutions simply by measuring the molecular weight of the ssDNA fragments. In the first approach, we assumed that a PCR reaction generated ssDNA copies of either the (+) or (-) strands, not both. It is possible to produce single-stranded copies by several techniques among which are (1) the use of only one primer (asymmetric amplification), (2) the use of two primers, one of which is end biotinylated, to provide for the selective removal of the strand it primes, and (3) the selective digestion of one of the strands by a exonuclease. For mass spectrometry to serve as an analytical tool for identifying the mutant forms, the 551+ and 553+ fragments have to be distinguished from WT+ or the 551- and 553- fragments must be distinguished from the WT-. This approach requires relatively high mass accuracy because the mass of each of the 6 oligonucleotides has to be determined with an accuracy having a deviation smaller than about ± 8 Da so that the (-) strands are distinguishable (See Fig. 1). The WT- and 553- differ by 16 Da so if the accuracy for mass determination has a deviation $< \pm 8$ Da, these strands can be distinguished theoretically. DNA samples isolated from CF patients will not contain a mixture of the WT and the mutant forms unless the patient is heterozygous in this part of their genome. We are not aware of the frequency of heterozygosity for the G551D and R553X mutations in the CF gene.

In the second approach we anticipated that the (+) and (-) strands will be present in a PCR amplified sample and that the differences in molecular weight of the (+) and (-) strands might be used to distinguish the presence of the WT or mutations. Without prior removal of one of the complementary strands produced during PCR amplification, as mentioned above, we could expect to see a mass spectrum that shows a pair of peaks indicating the presence of the WT+/WT-, 551+/551- or the 553+/553- fragments, providing the complementary strands can be denatured and prevented from annealing while the MALDI sample is prepared. The measurement of the difference in molecular weight of the complementary strands relaxes the mass accuracy requirements, noted above for the first approach, but necessitates high resolution.

The 553+/553- pair (DM =21.2 Da) will be separated with baseline resolution if the mass spectrometric technique (instrumental + sample preparation factors combined) shows a resolving power (M/Dm) of ~300. A mass spectrometer operating at a resolution of 300 will also resolve the WT+/WT- pair (DM= 52.2 Da), and the 551+/551- pair (DM=83.2 Da) will also be resolved because their mass differences are greater. We have recently demonstrated MALDI-TOF-MS resolution approaching 1000 for oligonucleotides. Therefore, it seems reasonable to expect to discern the differences in molecular weight (MW) of the pairs. The relative size of the fragments that we analyzed is shown in Fig. 1.

Methods

We synthesized the six oligonucleotides, listed above, with solid-phase-based phosphoramidite chemical procedures incorporated into a new 12-channel automated format.[8] The oligonucleotides were synthesized with 5'-terminal trityl groups which were later removed during column purification or were synthesized without a terminal trityl group and HPLC purified. We described previously the linear MALDI-TOF mass spectrometer we used in this investigation. In summary, we use a pulsed N_2 laser (λ_{em} = 335 nm) to desorb analyte ions (-30 kV acceleration voltage) from a 3-hydroxypicolinic acid (3HPA) matrix and time their flight through an approximately 1 m. long flight tube. The laser intensity was adjusted to a level that was slightly greater than threshold for production of analyte ions, ranging between 30 and 80 mJ/cm^2 . Not every laser pulse produced an analyte ion at the slightly greater than threshold intensity. TOF spectra were summed and averaged and this information is included in the figure captions of the mass spectra. Mass spectrometer samples were prepared by evaporating mixtures of 3HPA, analyte, and diammonium citrate, on the MALDI sample holder. The samples were typically prepared by evaporating, with blowing N_2, the mixture of 1 µL of saturated 3HPA in 1:1 acetonitrile-water (0.5 M), 1 µL of 10^{-4} M DNA in water, and 1–2 mL of 0.1 M aqueous diammonium citrate. The surface of the mass spectrometer sample holder, roughened with 400 grit sandpaper, facilitated small crystal formation, and the process of rubbing the surface during sample evaporation with the tip of a disposable micropipette tip significantly enhanced mass resolution. Some of the samples were prepared by the method of Vorm *et al.*[9] but we found this procedure less satisfactory. When we report molecular ratios of matrix/analyte, it indicates our use of the former technique.

Results and Discussion

We analyzed the individual oligonucleotides with and without internal calibration standards (see Figs. 1–6). One of two additional oligonucleotides, a 10-mer (3027 Da) or a 15-mer (4569 Da) and the molecular ion of the 3HPA matrix served as the internal calibrants. When internal calibrants were not utilized, we used a calibration factor obtained from the previous analysis of a number of other DNA samples. Bold format in Table 1 indicates the results from the analyses with

internal calibrants while non-bold format identifies results obtained without internal calibrants. Suffixes (a) and (b) refer to two different synthesis batches for the oligonucleotides.

Table 1. Measured masses of individual oligonucleotides.

oligonucleotide	calc. MW, Da	meas. MW, Da	deviation, Da
WT+(a)	6873.4	6864, **6870.2**	- 9.4, **- 3.2**
WT+(b)	"	**6870.3**	**- 3.1**
551+	6857.4	6847, **6853.4**	-10.4, **- 4.0**
553+	6888.4	**6886.9, 6881.8**	**- 1.6, -6.6**
WT-(a)	6925.6	6919, **6920.3**	- 6.6, **- 5.3**
WT-(b)	"	**6922.7**	**- 2.9**
551-	6940.6	6932, **6933.7**	- 8.6, **-6.9**
553-	6909.6	**6908.7, 6907.4**	**-0.9, -2.2**

The 10 samples in Table 1 analyzed with internal calibrants show an average deviation of -3.7 Da from the expected mass. With the 2-point calibration we applied to the mass data, all but one of the measured masses correctly identified the unique oligonucleotide. The 6933.7 value for the 551- fragment is 6.9 Da from its calculated mass and 8.1 Da from the mass of the WT- fragment, yielding a measured mass nearly equidistant from each. The assignment of this mass value to the correct oligonucleotide is questionable. The systematic determination of masses less than expected (-3.7 Da) attributes to a small time delay for the start signal to reach the digitizer, introducing an offset in the start time. Knowledge of the systematic deviation, however, allows one to normalize the data, and then identify each oligonucleotide correctly.

The spectra in Figs. 2–7 show additional features worthy of comment. All of the oligonucleotide spectra had a predominating component at the expected mass except 551+ (Fig. 3) which had an additional component with a mass of 7155 Da. This component is 303 Da greater in mass than expected, which might correspond to a molecule with an additional thymidine. An additional thymidine adds 302 Da but a preferable assignment for the 7155 Da peak is the 22-mer with a 5'-dimethoxytrityl group conserved because of incomplete deprotection at the end of the oligonucleotide synthesis process. Hathaway[10] reported a similar observation. If the oligonucleotide synthesizer were programmed improperly so as to synthesize a 23-mer instead of the 22-mer (551+), then this 22-mer should not be observed in the spectrum.

The spectra of the individual oligonucleotides show a resolution for the oligonucleotides which ranges from 470 to 810. Small peaks at masses smaller than the predominating oligonucleotide peaks indicate the presence of minor contaminants. The 3027 Da calibrant was not purified and shows several associated contaminants having masses a few hundred Da less than the calibrant. The 553 +/- oligonucleotides were HPLC purified and show about the same contaminant levels as the other oligonucleotides which had been column purified. Chromatograms recorded during the HPLC purification of the oligonucleotides indicated only the presence of contaminants having a molecular weight much smaller than the expected oligonucleotide. These spectra (Figs. 1-6) demonstrate that the mass spectrometric analysis of synthetic oligonucleotides provides a way to assess quality control for oligonucleotide synthesis.

When complementary oligonucleotides were mixed together during MALDI sample preparation, they annealed and could not be detected in the double-stranded form. If, however, we heated the annealed oligonucleotides to 100°C for 1 minute and then cooled them rapidly, we preserved the denatured oligonucleotides in the dry MALDI sample and observed them in mass spectra. Simply by placing 1-2 μL of the heated oligonucleotide onto the sample probe, along with

several μL of room temperature matrix solution, the evaporative cooling created by a jet of N_2 used to evaporate the sample prevented reannealing of the complementary strands. The spectra of the denatured oligonucleotide pairs are shown in Figs. 7–9 and the results are summarized in Table 2. We produced baseline separation for the 551+/- pair and resolved the other two pairs, but not as well. Each oligonucleotide pair was resolved adequately to allow correct allele identification and the results demonstrate a new way to identify base-pair substitutions that does not rely on gel electrophoresis or sequencing reactions. Gel electrophoresis routinely separates 23-mers from 22-mers but the gel-running time extends to about an hour compared to a few minutes for mass spectrometric analysis. The mixed oligonucleotide samples showed a resolution as high as 920 (Fig. 9).

Table 2. Complementary oligonucleotide analysis following denaturation with heat.

fragment	expected, Da	fragment difference	measured, Da	difference	correct assignment
WT+	6873.4		6873.9		
		52.2		51.5	Y
WT-	6925.6		6925.4		
551+	6857.4		6865.3		
		83.2		79.9	Y
551-	6940.6		6945.2		
553+	6888.4		6882.9		
		21.2		21.1	Y
553-	6909.6		6904.0		

We created dsDNA by annealing a 22-mer with a complementary 23-mer. These oligonucleotide pairs had small molecular weight differences. If the (+) and (-) forms of the alleles we created contained the same number of bases, the difference in mass of the ssDNA strands would have been even larger, thus relaxing the required analytical resolution. In this regard, the 553+/- pair differs in molecular weight by 21.1 Da even though the 553- strand is one base longer than its complement. The removal of one base from the 553- strand would decrease its molecular weight by 289 Da if the cytosine on the 3' end were deleted. In so doing, the difference in weight between the 553+ and 553- strands increases to 268 Da and makes the analysis easier with respect to resolution considerations. Our decision to study the separation of 22- and 23-mers derived from the choice of endonucleases that could be utilized to isolate the selected region of the CF gene. The extension of our approach to identify base substitutions in complementary strands longer than those reported here by measuring the difference in molecular weight of the denatured strands appears reasonable but depends on the sequences and their attendant molecular weight differences. The selection of alternative endonucleases might resolve the difficulties by providing ways to generate double-stranded fragments comprising single strands that have a larger difference in molecular weight. The mass spectrometric approach described here eliminates the time-consuming electrophoresis separations normally applied to PCR products, making analysis by MALDI-TOF-MS alternatively attractive.

Conclusions

We demonstrated the capability to mass analyze denatured dsDNA as large as 23 base pairs (~6900 Da) with a mass resolution (M/Dm) as high as 920 and a mass accuracy deviation of -3.7 Da for samples prepared with internal calibrants. We applied MALDI-TOF-MS to the analysis of short ssDNA by denaturing the dsDNA with heat before the matrix and oligonucleotide crystallize

during sample preparation. With this technique, we identified base substitutions in 22 base-long dsDNA fragments by comparing the difference in molecular weight of the complementary strands of wild type alleles to mutated alleles, a procedure that does not require sequencing. A mass resolution for oligonucleotides approaching 1000, as we report, provides a way to analyze complementary strands of DNA at least as long as 22 base-pairs. With longer asymmetric overhangs, the difference in molecular weight of the complementary strands increases, thus extending our approach to larger fragments of dsDNA. This analytical approach should be applicable to PCR generated samples, providing they are purified to remove contaminants normally present in a PCR cocktail.

Acknowledgments/Disclaimer

We thank Dr. K.J. Wu for supplying the 10-mer and providing a procedure for utilizing 3HPA and diammonium citrate. Prof. J. Rine suggested the application of MALDI-MS to identify the base substitutions and Dr. A. Lishanskaya selected the region of interest in the CF gene.

This work was supported by the Director, Office of Energy Research, Office of Biological and Environmental Research, Analytical Technology Division, of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

References

1. Arnheim, N., H.H. Li, and X.F. Cui (1990) Genomics, **8**, 415-419.
2. Zielenski. J., Fujiwara, Markiewicz, D., Paradis, A.J. (1993) Am. J. Human Genetics, **52**, 609-615.
3. Griesen, K., Loeffelholz, M., Purohit, A., and Leong, D. (1994) J. Clinical Microbiol., **32**, 335-351.
4. Rotbart, H.A. (1991) Clinical Microbiol. Rev., **4**, 156-168.
5. Schneider, K., Chait, B. (1993) Org. Mass Spectrom., **28** 1353-1361.
6. Pieles, U., Zurcher, W., Schar, M. and Moser, H.E. (1993) Nucleic Acids Res., **14**, 3919-3196.
7. Benner, W.H., Horn, D., Katz, J., Jaklevic, J. (1995) submitted to Rapid Commun. Mass Spectrom.
8. Sindelar, L., Jaklevic, J. (1995), submitted to Nucleic Acids Res.
9. Vorm, O.P., Roepstroff and Mann, M. (1994) **66**, 3281-3287.
10. Hathway, G.M. (1994) **17**, 150-155.

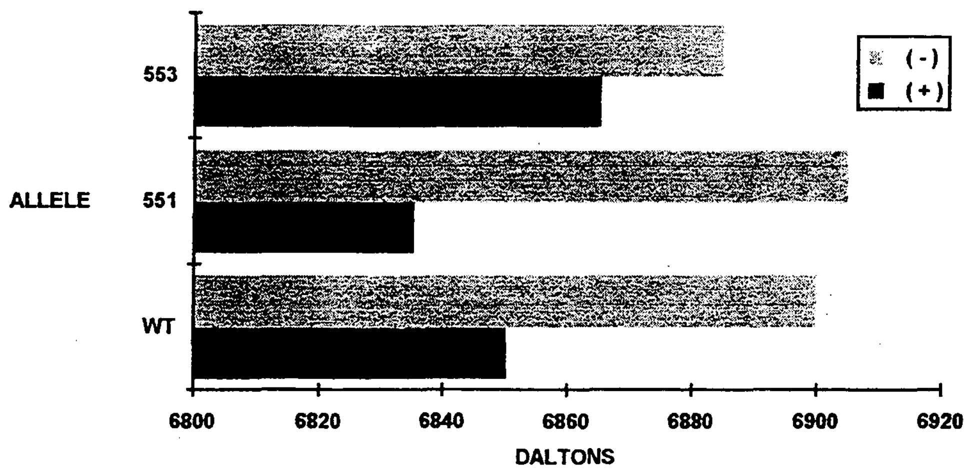


Figure 1. A comparison of the lengths of the six DNA fragments we attempted to identify by MALDI-TOF-MS.

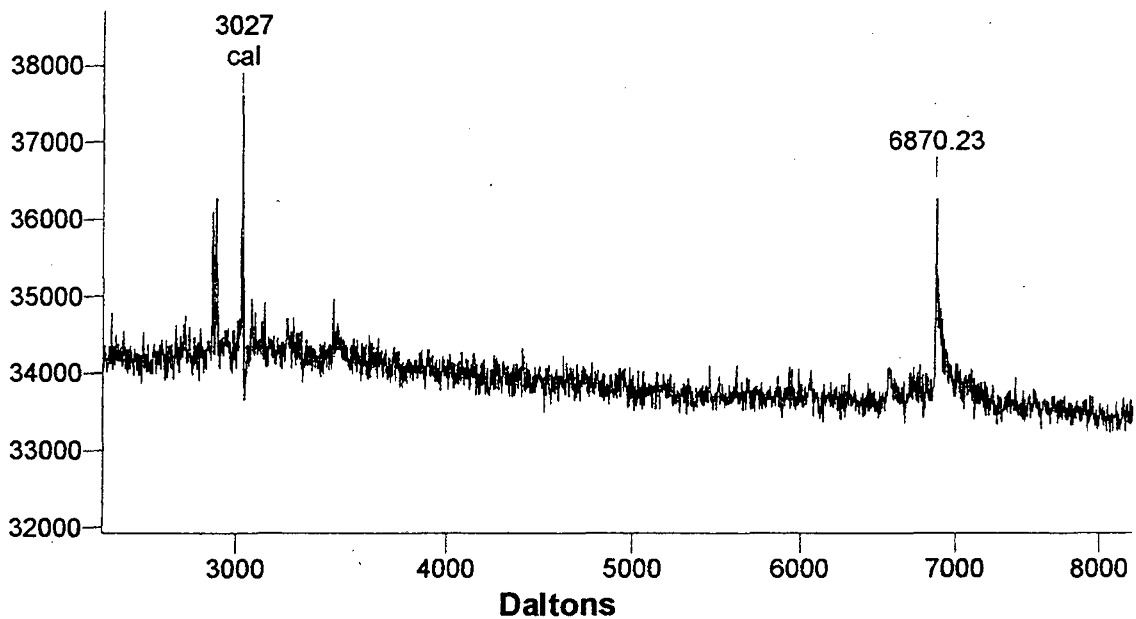


Fig. 2. Mass spectrum of WT+ in 3HPA matrix, Mr = 5000, 582 shots averaged, -25 kV, resolution of the 6870 Da peak is 810. (cfg52.mas)

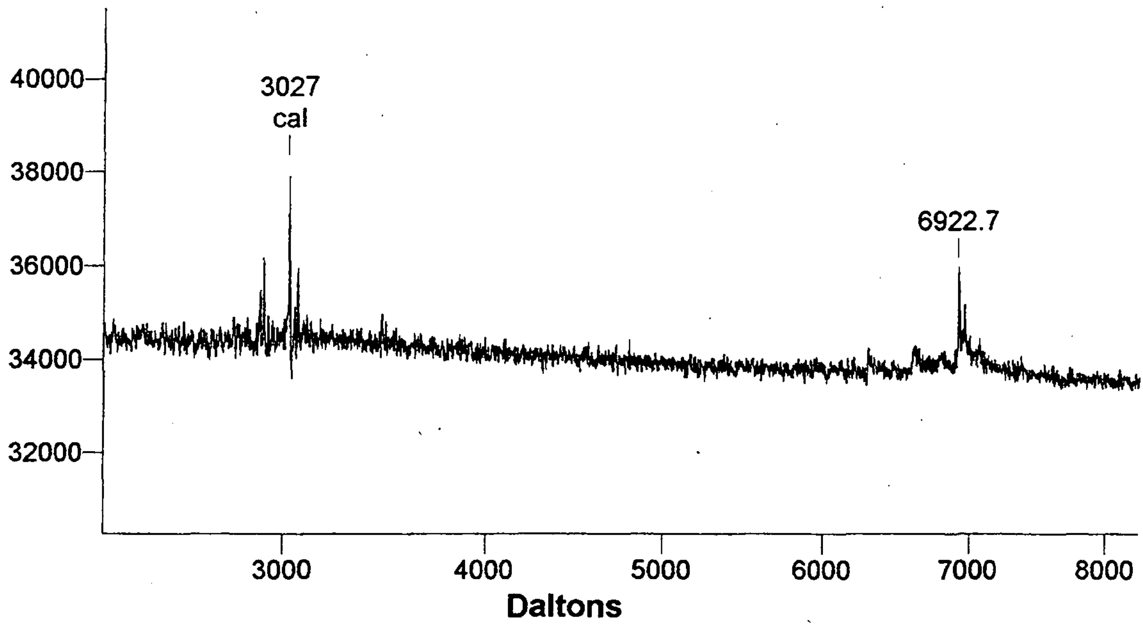


Fig. 3. Mass spectrum of WT- in 3HPA matrix. Mr= 5000, 621 shots averaged, -25 kV, resolution of the 6922 Da peak is 520. (cfg81.mas)

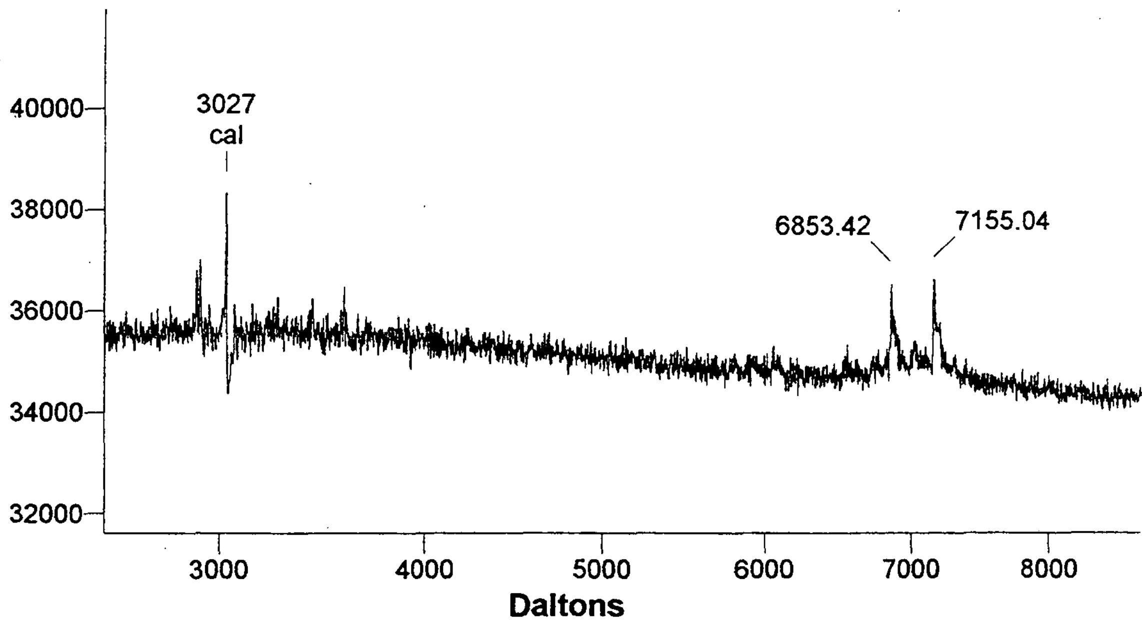


Fig. 4. Mass spectrum of 551+ in 3HPA matrix. Mr = 5000, 598 shots averaged, -25 kV, resolution of the 7155 Da peak is 300. (cfg92.mas)

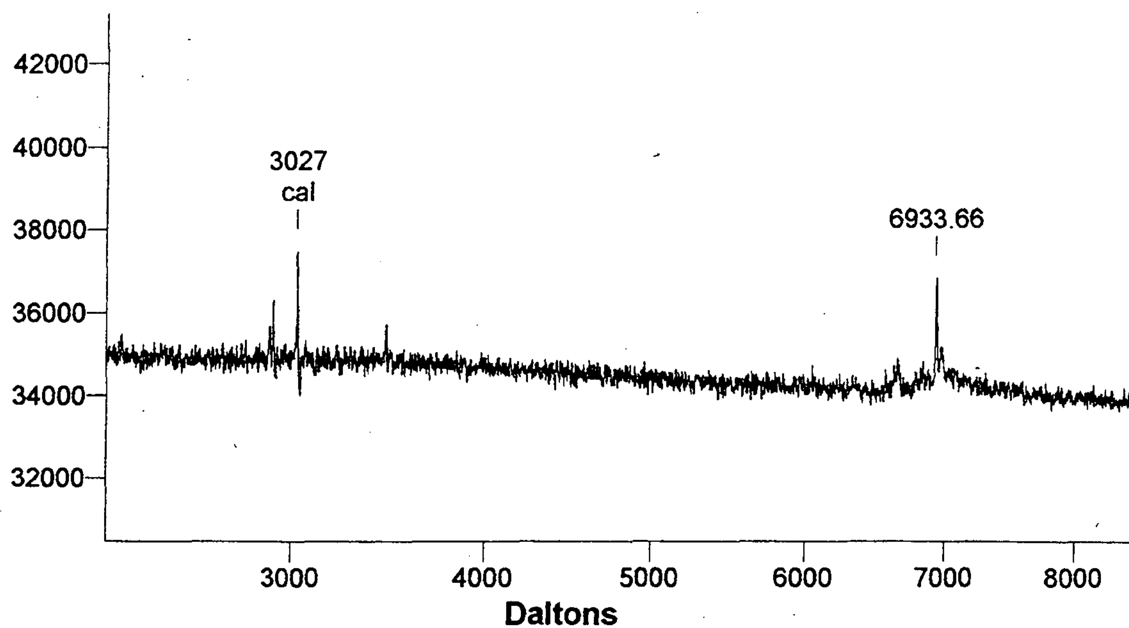


Fig. 5. Mass spectrum of 551- in 3HPA matrix. Mr = 5000, 584 shots averaged, -25 kV, resolution of the 6933 Da peak is 580. (cfg102.mas)

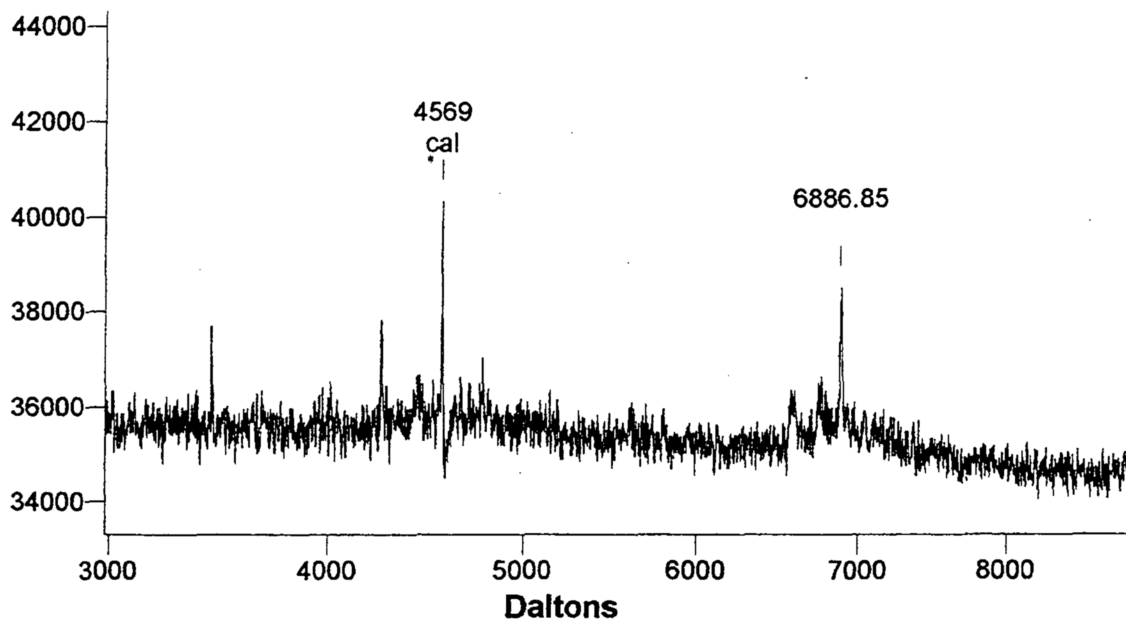


Fig. 6. Mass spectrum of 553+ in 3HPA matrix. Mr = 1500, 573 shots averaged, -25 kV, resolution of the 6886 Da peak is 540. (cfgb4cal.mas)

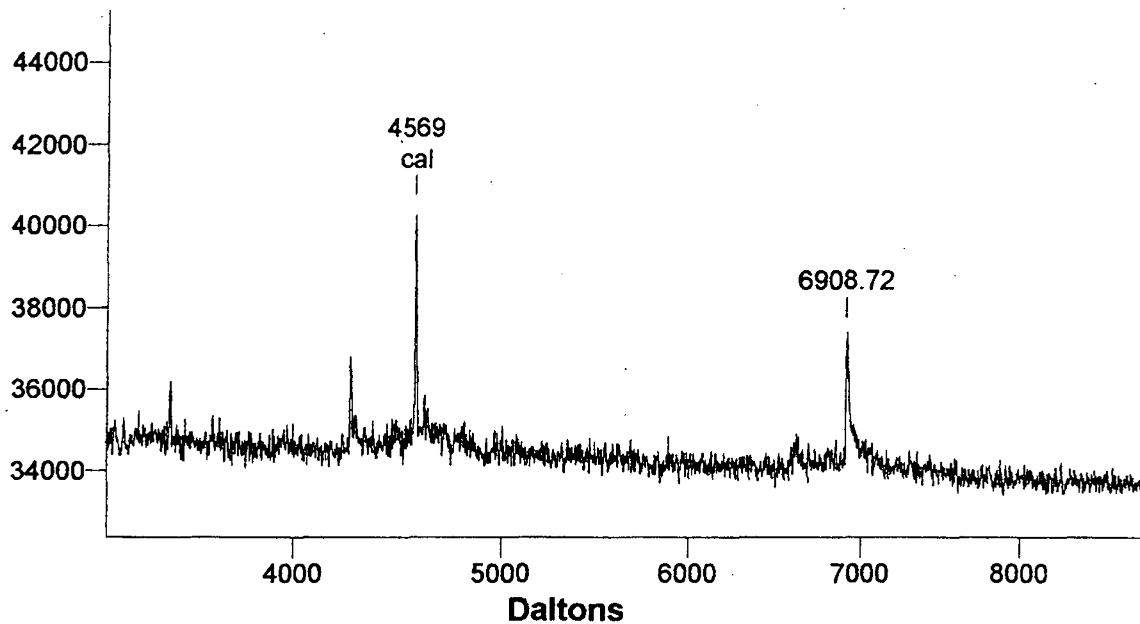


Fig. 7. Mass spectrum of 553- in 3HPA matrix. Mr = 1500, 725 shots averaged, -25 kV, resolution of the 6908 Da peak is 470. (cfgb5cal.mas)

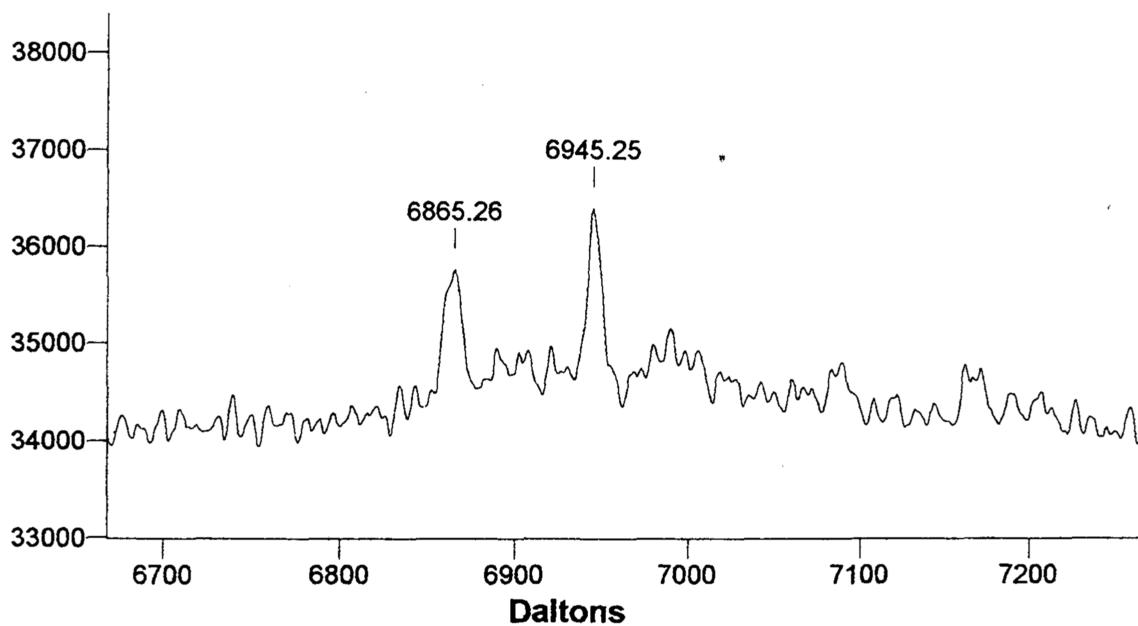


Fig. 8. Separation of 551+/- complimentary strands in 3HPA matrix. Mr = 5000, 684 shots averaged, -25 kV, resolution of the 6865 Da and the 6945 Da peaks are 525 and 700, respectively. (cfg9_10c.mas)

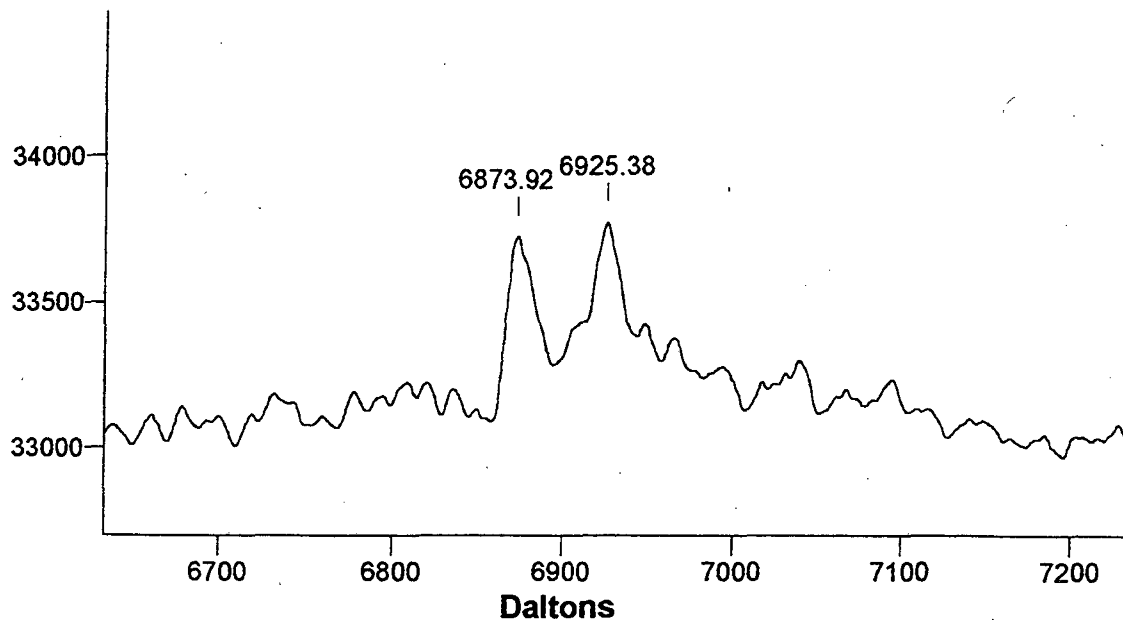


Fig. 9. Separation of WT+/- complimentary strands in 3HPA matrix. Mr = 3000, 500 shots averaged, -25 kV, resolution of the 6873 Da and the 6925 Da peaks are 330 and 340, respectively. (cfg5_6a.mas)

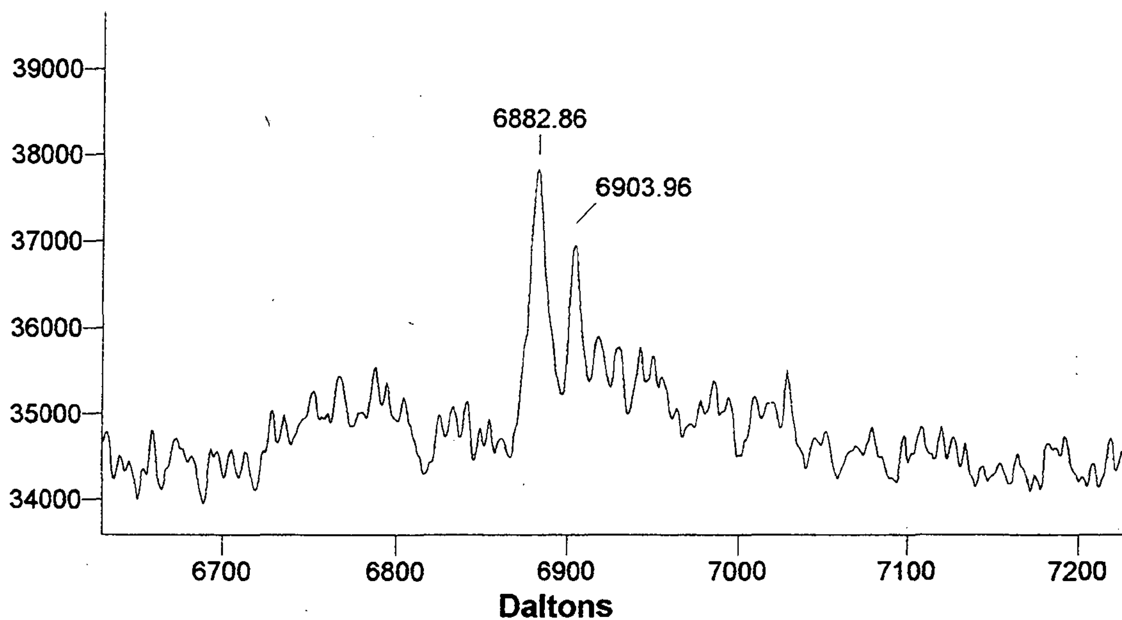


Fig. 10. Separation of 553+/- complimentary strands in 3HPA matrix, Mr = 1000, -25 kV, 665 shots averaged, resolution of the 6882 Da and 6903 Da peaks are 920 and 750, respectively. (cfgb4_5.mas)

LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
TECHNICAL AND ELECTRONIC
INFORMATION DEPARTMENT
BERKELEY, CALIFORNIA 94720