Contents.

1. Introduction.

The DNA sequences of genes and genomes are the basis for advance in understanding of life, molecular biology, medicine, biotechnology and studies of the evolutionary process. The challenge of unravelling the complexities of genomes has advanced the development of new strategies for genome analysis. The progress in multiple genome research programs is dependent on improved conventional methods, such as cycle sequencing (Sanger et al., 1977) and novel strategies, complementary to the old methods. The Sequencing by Hybridisation methodology proved to be such a technology which considerably assists the task of complex genome analysis. To date this technique has proven to be useful for a wide range of genetic analysis applications which include DNA identification in diagnostics, DNA mapping and partial or complete sequencing of genes.

1.1 Sequencing by hybridisation

The concept of sequencing by hybridisation (SBH) has been independently proposed by several research groups within the last decade(Drmanac, 1987), (Southern, 1988), (Bains and Smith 1989), (Lysov, 1988). It utilises the basic principle underlying the functioning of living matter that is, each strand of DNA recognises a uniquely complementary sequence in a highly parallel manner through chemical specificity of base pairing. This process is termed ”hybridisation”. Any linear sequence can be regarded as an assembly of overlapping shorter fragments. The complete sequence of the target can be deduced from independent hybridisations with the full set of $\mathbf{4^n}$ oligonucleotide probes of length **n** consisting of 4 nucleobases (A, T, G, C)under specific conditions. The sequence reconstruction is done by assembling according to (n-1) overlaps of hybridised sequences with various computer algorithms.

Figure 1. The reconstruction of a sequence fragment with overlapping 8-mer oligonucleotides.

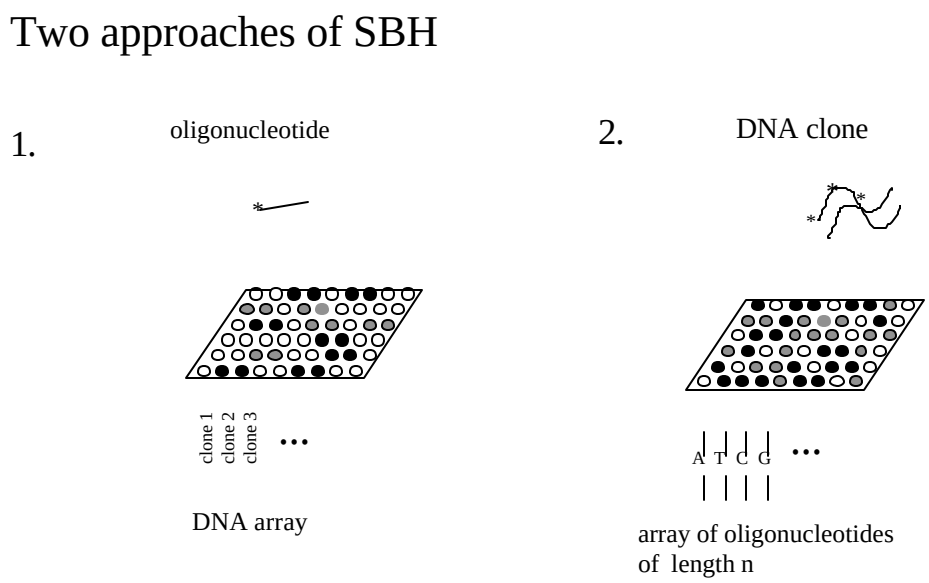
```
AGGGAGTG
 GGGAGTGC
  GGAGTGCA
-----------
5'--AGGGAGTGCA---3'
```

Octamers were considered to be probes of a compromised length. On one hand the probes had to be reasonably short so that the total number of all possible combinations was not indefinite, but not too short, so that the target can be unambiguously recovered from hybridisations. The full probe set thus would consist of 65,536 oligonucleotides representing all possible octamers (4^8). This number can be reduced by a factor of two by consideration of complementary oligonucleotide sequences (e.g. AGAGAGAG is redundant to TCTCTCTC because they are complementary) considering that assayed DNA is double stranded. Then recent developments

in sequence reconstruction algorithms allowed the required number of probes to be reduced further. For example, instead of 4^7 heptamers (16,384) only 3000 appeared to be sufficient for accurate complete sequencing of 1.1 Kbp (Drmanac et al., 1993).

The oligonucleotides which are 7- or 8-nt long are postulated in SBH to be more suitable to synthesise the complete set of probes and have the advantage of theoretically better discrimination of mismatches in hybridisation, i.e. are more specific than longer oligonucleotides. However, with decreased length of the probes the stability of duplex, also decreases. A number of methods were proposed to enhance the stability of such short probes: introducing chemically modified nucleotides (Hoheisel et al., 1990), or universal bases (Parinov et al., 1996) or degenerate bases at 5' and 3' ends (Drmanac et al., 1992)

Two basic fashions of the method were being developed in parallel. Both utilised arraying of either of the components of reaction between query DNA and oligonucleotide probes. In so called *format 1* DNA clones attached to solid support are hybridised with set of synthetic labelled oligonucleotides (Drmanac et al., 1996). In *format2* (Southern et al., 1992), (Khrapko et al., 1989) a complete set of oligonucleotides is arrayed on a surface and hybridised to an individual labelled target.



The second, so called *format 2* approach, is theoretically more informative since one experiment should provide full sequence information for given DNA clone. The method requires, however, the manufacturing of high quality arrays.

Initially, idea of SBH seemed to potentially enable faster determination of sequence. The technical barriers and limitations of the *de nuovo* sequence determination method became more obvious later.

First of all, it is difficult to achieve absolute specificity of hybridisation of a short oligonucleotide to the complex DNA, especially when multiple heterogeneous oligonucleotides are hybridised in parallel as in *format* 2.

The ambiguity is introduced by additive contribution of several mismatched basepairs of an imperfect match. Hybridisation is a complex reaction where various interactions between DNA nucleobases contribute to the formation of a duplex and to its stability. These include number of perfectly matched (fully complementary, according to Watson and Crick) and mismatched pairings, contribution of "dangling ends"(Williams et al., 1994) and the nearest neighbour context (Wetmur, 1991). In experiments efficient discrimination of the mismatches can be improved by tuning experimental conditions such as appropriate probe concentration and time and temperature of washing (Drmanac et al., 1990).

The other problem in the process of sequence reconstruction is frequently distributed in genomic DNA the low complexity regions such as polyA stretches or tandem repeats which create branching points leading to errors in reconstruction. Computer simulations have shown that complete set of 8-mers would allow the elucidation of 80% of random DNA sequences of 200 bases length (Khrapko et al., 1989).

1.2 Applications of SBH

The above challenges of the original method motivated researchers to focus on simpler but nevertheless powerful applications of SBH, on comparative sequence analysis rather than *de nuovo* sequence determination.

The viability of the approach was shown most convincingly for mutation detection and comparative sequence analysis in *format2* and large scale mapping and partial sequencing for *format1*. Specially designed oligonucleotide arrays containing all possible deviations within a known sequence were successfully applied to sequence polymorphism determination. Oligonucleotide chips comprising 7 or 8-mer sequences can be used to determine polymorphisms in a single transcript of a few kilobases in length (Drobyshev et al., 1997), and much more complex arrays were used to identify polymorphisms of the mitochondrial DNA (15-mer array, (Chee et al., 1996) or to perform highly parallel phenotypic analysis of the entire organism of yeast (array of 20-mers (Lockhart David et al., 1996). Such applications however, require the use of a reference control of known sequence for quantification, so that any difference in signals between the reference and query sample indicates mutation (Chee et al., 1996) or difference between organisms (Lockhart David et al., 1996).

1.3 Partial SBH

Biologically relevant information with partial SBH can be assessed using an incomplete list of oligonucleotide probes.

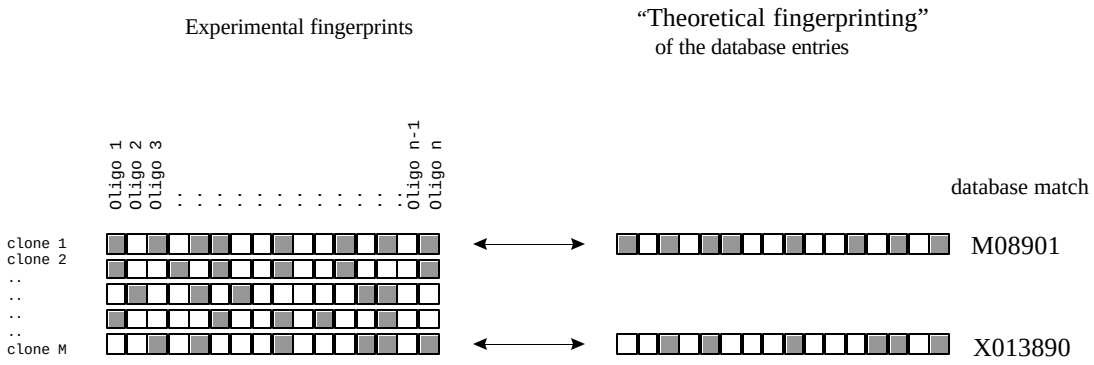
Initially this strategy was proposed as method to generate high resolution maps of genomic fragments (Poustka 1990). The application of this approach was tested on model cosmid library using pre-selected 12-mer oligonucleotides which preferably hybridised to the insert DNA (Craig et al., 1990). Later an approach using non-

unique shorter probes which had higher frequency of hybridisation was proposed (Lehrach et al, 1990). It was shown that the number of such non-unique probes needed for a library characterisation is genome-size independent (Hoheisel et al., 1991).

1.3.1 cDNA library catalogues

The partial SBH of randomly selected cDNA clones was proposed as a method of identification of active genes and simultaneous determination of their transcription level in tissue (Drmanac et al., 1996) or organism (Meier-Ewert et al., 1998), (Poustka et al., 1999) specific libraries by determining the abundance of clones derived from the same transcript.

The sequence signatures (oligofingerprints) obtained with non-exhaustive probe sets can be compared to the fingerprints derived from sequences in the database. Such comparison has the potential to assign sequence and even function to a query sequence if any functional motifs are revealed in the signatures. The strategy relies on common occurrences of distinct probes in query and known subjects from database.



Extensive application of fingerprinting- based analysis should assist in the completion of genome sequencing by comparison with already sequenced genomes of similar organisms("standards") based on hybridisation data alone.

The ability to perform simultaneous analysis of many clones in parallel allows the extension of the fingerprinting approach to characterise gene expression by analysing the cDNA libraries. The analysis of cDNA libraries potentially leads directly to gene discovery, since these libraries represent only expressed gene information. In this way fingerprinting of an infant brain library revealed 20,000 distinct genes (Drmanac et al.,

1996) and it was estimated that a similar number are expressed but not represented in a given collection of cDNA clones.

Using sequence signatures or fingerprints for cDNA characterisation allows 1) the grouping (termed "clustering") of clones by comparison based on partial sequence data and thus effective normalisation of libraries; 2) quantitative expression studies of both known and still uncharacterised genes across different cell types or developmental stages using highly redundant non-normalised libraries. The fingerprints generated by oligonucleotide hybridisation can identify clones which are identical or show similarity to the known genes, or represent the new genes. The drawback of the method is that representation of some mRNAs are distorted during cloning process and might not be estimated correctly and may not provide accurate measure of these transcript's level in vivo.

1.3.2 Shotgun sequencing

Apart the above applications the strategy of grouping by fingerprints was extended for minimisation of redundancy of shotgun sequencing. When larger genome fragments are placed on the map, detailed sequence determination includes further sub-fractionation into pieces suitable for direct sequencing which is length-limited. The libraries of such sub-fragments are termed shotgun libraries.

In shotgun cloning efficiency of different fragments is inevitably variable. Obtaining minimal representation in sequencing of each region of query sequence requires redundancy of sequencing effort since the relative position of each shotgun clone is unknown. As a result in random picking of clones for sequencing in a typical 5-fold or 6-fold redundancy shotgun sequencing project, some regions inevitably have 10-fold coverage (Demolis et al., 1995). Practically for a given cosmid of 100 kb length an average of 11 fold coverage up to 30-fold coverage in some regions is typical for conventional shotgun sequencing library (Radelof, personal communication).

Pre-selection of clones can reduce sequencing effort. The generation of oligofingerprints for the shotgun clones prior to sequencing allows the sorting of clones based on their partial sequence data. A minimal overlapping set of clones can be generated by selecting clones with the most dissimilar fingerprints following pair-wise comparison. In the final stage of the data processing in order to close gaps fingerprinting data is used to identify clones with similarity to chosen in the first round.

This effort is justified by the reduction of sequencing effort by a factor of two (Radelof et al., 1998) up to factor of 5 in a simulated 3Mb genome (Drmanac et al., 1996). Simultaneously finer mapping of gene-specific sequences, such as promoter element- or CpG island with specific probes could provide sequence-based information for further analysis.

1.4 Data analysis: scoring functions and database matching

Because of non-specific oligonucleotide hybridisation the result in partial SBH applications is largely dependent on reliable analysis of data. To assure the reliability of a clustering algorithm it is essential to have reliable controls for clustering and data processing.

Clustering uses the composite value from hybridisation scores which has been calculated similarly by different authors (Milosavljevic et al., 1995), (Herwig et al., 1999). A core of these methods hybridisation is the ranking of intensity values for a given oligonucleotide hybridisation or of several hybridisations to a single target clone. The rank scores of 1 and 100 are assigned to the probes with the smallest and the highest value, and the ranks for all the rest oligonucleotides which hybridised to a clone are uniformly distributed in between.(Milosavljevic et al., 1995) Such scaling leads to a enriched information extracted from a fingerprint than from values transformed to 0/1 (minus or plus) signal (Milosavljevic et al., 1995).

Further analysis implies using the fingerprinting data obtained by hybridisation with 100-200 oligomers for database searches. Ideally the recognition of genes which gave rise to clusters should be done by comparing with entries extracted from databases upon blasting the oligomer list against the databases, i.e. experimental and theoretical fingerprints. The match is assigned to clones sharing the signature and is verified by direct sequencing. One should be aware that resolution of fingerprinting might result in identification of genes only showing highly significant sequence similarities: in (Meier-Ewert et al., 1998) the comparison has shown that only from clusters with significant database matches ($P<10^{-7}$) 90% were confirmed by sequencing and BLAST (Altschul et al., 1990) search. In searching for new genes only clones with unique fingerprints are processed.

In a genome-scale experiment such sequence determination by hybridisation and database matching with no preliminary sequence reconstruction was performed using a *E.coli* genomic library (Milosavljevic et al., 1996). Firstly the experimentally obtained hybridisation with nearly 1000 oligonucleotide probes about 1/10 of positive oligonucleotides (and their complements) based on criteria of the highest scores among all positive probes were selected to form a fingerprint. The theoretical signatures based on complementarity of the probe set to the sequences in the database were compared to the experimental. In this study 60% of clones were identified correctly by comparing experimental and theoretical signatures as then was confirmed by sequencing.

The reason for that can be high level of non-specific hybridisation in non-stringent conditions. This overestimates the complexity of the library and leads to false clustering, such as splitting a single cluster into many clusters or else erroneous co-clustering of the sequences.

1.5 The libraries used in oligonucleotide fingerprinting

While designing an experiment with multiple hybridisations it is important to take into account the source of DNA target. The versatility of fingerprinting methods is reflected in its application to various in source DNA libraries.

For example shotgun libraries are usually size-selected to 1-2kb of the clone length to ensure full coverage by one sequencing read and have differing degrees of overlap. The characteristic difference of cDNA clones from genomic clones is that cDNA derived from same mRNA species have a shared 3'-end (when primed with poly dT primers (Clark et al., 1999) which results in high degree of uniformity in hybridisation behaviour and provides a basis to group together cDNAs derived from the same transcript even though they have different length (500 bp to 5000 bp). This heterogeneity in length however results in different hybridisation frequency of oligonucleotides within the cDNA clones.

Naturally occurring or selected pools of shorter mRNAs as well as genomic clones of restricted length require an increased number of probes to be used in analysis because the frequency of a given oligomer in shorter clones is lower. A possible solution for increasing hybridisation frequency is the use of pools of oligomers or oligonucleotides with degenerated internal position or the shorter probes.

The complexity of a library has also implications for the amount of data necessary for library characterisation. Unlike libraries prepared from the whole genome when the initial representation of different regions is equal, in the cell the abundance of certain RNAs is different thus for a qualitative analysis the prepared library size should be sufficiently large to ensure adequate representation of as many RNA species as possible. The complexity of RNA population present in the cell has been estimated by various authors with all estimates approximates to 30,000 genes in complex brain tissue (Drmanac et al., 1996). Such complexity estimation dictates that the representative library size would approach 100,000 -1,000,000 clones libraries analysed with 200-300 octamer hybridisation probes (Drmanac and Drmanac, 1999).

The statistical studies indicate that oligonuclotide probe sets can be optimised for the sequence content of the genome ((Cuticchia et al.,)) or expressed sequences (Herwig, in preparation). It was shown that the distribution of all possible 8-mers over the real genomic DNA fragment had organism specific maximums of occurrence, such as 20 times in random DNA and 5 times for bacterial and yeast genomes (Mayraz and Shamir, 1998).

The theoretical rationale for the probe set design is the achievement of frequency of hybridisation so that partitioning of the analysed library is maximal. The optimum is achieved at 50% frequency of hybridisation (Clark et al., 1999). A lower than this estimate for the hybridisation frequency of random 8-mer oligonucleotide to random DNA library(approximately 3%, Herwig, in preparation) is confirmed by experimental data. Given the fact that DNA structure is biased - e.g. sub-word content in coding and non-coding fragments is different and that there are organism specific differences, the probe set can be specifically selected depending on the DNA source. Specifically-designed probe sets adopted to experimental material can outperform the totally random set by 20%(Schmitt A.O., personal communication), which results in a smaller number of hybridisations which are sufficient to characterise these libraries.

Designing of probe for the specific purpose of looking for certain genes is possible using algorithms which screen databases in order to select oligonucleotides of specified length with high frequency of occurrence in gene families or gene functional units, such as ribosomal gene promoters or exons or transcription factor binding sites and

low frequency in any other site (Kolchanov et al., 1995). Interestingly, such probes pre-selection for identification of functional homologues doesn't require homology of sequences themselves as was shown in *in silico* analysis of EST library containing 713,000 cDNA clones with oligonucleotide set specific for G-coupled receptor proteins(Kel et al., 1998). In this study 70% of G-protein receptors genes could be correctly identified *in silico* with 0.02% false positives.

1.6 New approaches and optimisation

Recently developed procedures for the automated processing of hundreds of thousands of clones: by replication, PCR-amplification and arraying (Drmanac et al., 1992), (Maier, 1995) enables the management of such large libraries. More than 10^7 individual clone/probe hybridisations can be performed routinely (Milosavljevic et al., 1996).

In order to increase the experimental data flow implied development of high throughput procedures was required. In oligofingerprinting considerable scaling up of the library handling processes were made in the automated picking, handling, replicating and gridding of the libraries (Maier, 1995). A huge part of the progress was subsequently due to automated raw data analysis and development of reliable algorithms for processing of the hybridisation scores. The further improvements in the technique should be the use of non-radioactive labelling, multiplexing for example by simultaneous use of differently labelled probes and higher grid densities. Another requirement is the design of oligonucleotide probes with a higher frequency- and sequence-specific hybridisation.

1.7 Non-radioactive labelling and detection methods

In the high throughput analysis the used of alternative to hazardous radioactive detection methods is desirable. In systematic hybridisation assays the criteria for choice of the label are the following: speed, simplicity, sensitivity and specificity. The recent progress in chemical technologies allows the incorporation of various fluorescent dyes into oligonucleotides via synthesis. Fluorophores covering the full spectrum of wavelengths are now available (Rhodomine, Fluoresceine, Cy-dyes etc. (Haugland, 1996)). Application of these dye in detection requires the use of glass as DNA immobilisation support which does not contain enough DNA material and can be only used once (Schena et al., 1996). High background on the nylon membrane as a DNA support is the main obstacle to the use of fluorescently labelled oligonucleotide probes.

The best alternative to the low molecular fluorophores is the fluorescent protein Phycoerythrin as can be judged from its fluorescent characteristics (Table). It has been used successfully as secondary detection

reagent conjugated to streptavidin for detection of hybridisation of biotinylated DNA or RNA on glass (Wodicka et al., 1997).

Fluorophore	**Absorption, nm**	**Emission, nm**	**Extinction coefficient (M^{-1} cm^{-1})**	**Quantum yield**	**Molecular weight, Da**
Phycoerythrin*	498, 545	578	$1{,}96 \times 10^{6}$	0.84	240,000
CryptoFluor^{tm} Crimson *	585,625	660	$4{,}6 \times 10^{5}$	n/a	40,200
PBXL-1*	545	666	n/a	0.6	1500,000
Cy5	649	670	$2{,}5 \times 10^{5}$	>0.28	798
Fluorescein	495	525	$8{,}4 \times 10^{4}$	0.5	390

* Fluorescent proteins

The family of phycoerythrin proteins is found in cyanobacteria (Bryant and Cohen-Bazire, 1981) and red algae. R- phycoerythrin is the basis of commercially available fluorescent dye (Oi et al., 1982) and is a protein of at least 3 sub-units. The molecular weight was reported to be between 214,000 and 290,000Da. The number of fluorophores per molecule is ca 35 (Intergen company (Purchase, NY), personal communication)

An alternative to direct fluorescent dyes are indirect chemo- or luminescence, where the multi-step procedure involves hybridisation and then detection of hapten-labelled oligonucleotides with anti-hapten-enzyme conjugates (Durrant and Fowler, 1994). Several available enzymes, AP (Alkaline Phosphatase) and HP(horseradish peroxidase) allow flexibility in the choice of detection methods to be used, though of all reporter groups AP would be still the only one approaching autoradiography. The introduction of new chemiluminescent substrates affords significant improvement. There are currently 3 of the substrates (Attophos (Cherry et al., 1994), CSPD and CPD-Star) with detection limits greater, than radioactivity, that are achieved in one tenth of the time required for the radioactive method (Bronstein et al., 1994), (Tizard et al., 1990).

The covalent attachment of synthetic oligonucleotides to the enzyme was shown to attain high sensitivity and reduced background compared to the multi-step procedure. Unfortunately, the methods developed for cross-linking of proteins to DNA perform poorly on short oligonucleotides such as used in oligonucleotide fingerprinting. The majority of protocols include 5'-end modification of an oligonucleotide so

that it carries the thiol- or aminogroup in order to couple it with an activated enzyme. The main drawback of the existing schemes is that they all require intermediate purification steps removal of reagents and all the modifications do not have 100% yield, thus separation from unreacted oligonucleotide is often required. To date no high throughput labelling procedure was applied so far. For very short oligonucleotide probes no reliable routine purification methods are available.

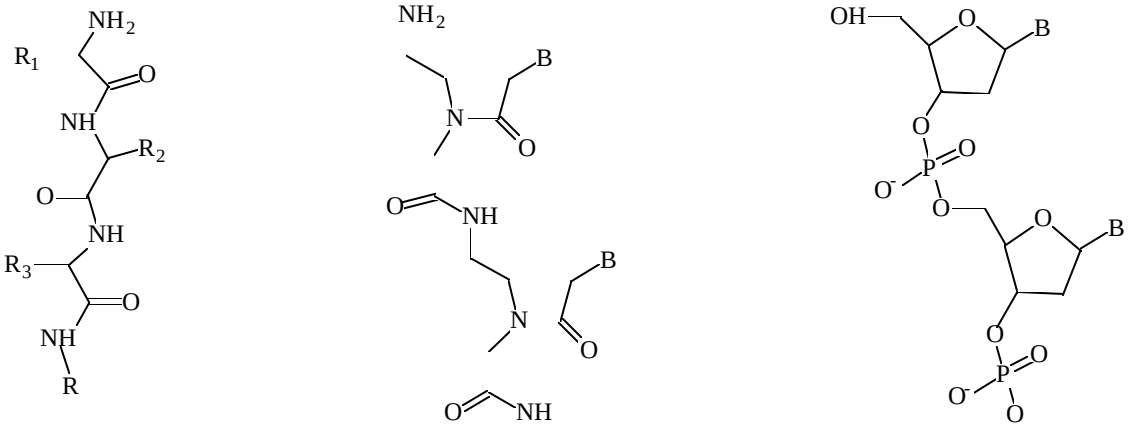
It was shown that enzymes can be coupled to short oligonucleotides via biotin-streptavidin link (Guerasimova et al., 1999). Hybridisation properties of such conjugates are not impaired by such modification. The method exhibits flexibility since many commercially available conjugates can be used depending on detection facilities. It is a cheap and automation alternative to radioactive labelling which enables automation. Unfortunately the AP schema can only be used for oligonucleotides larger than 9 nucleotides. The solution can be the use of different enzyme system or of more stable DNA analogues.

1.8 Chemical category of PNA

Oligonucleotide analogues are of major interest as a tool in molecular biology. In recent years many new DNA mimics with improved properties such as increased binding affinity to complementary nucleic acids have been introduced.

From other new compounds non-ionic oligomers are of special interest because of their phenomenal stability in complexes with DNA. In 1991 Nielsen (Nielsen et al., 1991) introduced a new class of such uncharged nucleic acid analogues known as PNA (peptide nucleic acid). In PNA the sugar-phosphate backbone is replaced by N-aminoethylglycine-based peptide-like structure and the bases are attached via a methylene-carbonyl linker. A free aminogroup at the 5'-end allows to the use of amino-active compounds for PNA modifications. The C-terminus of PNA corresponds to the 3' -end of DNA. PNA form duplexes according to Watson-Crick rules (Egholm et al., 1993).

peptide PNA DNA



It has been confirmed by independent studies (Tomac et al.,), (Jensen 1997) that PNA/DNA duplex has greater stability which is ascribed to the electro-neutrality of PNA. Compared to a given DNA/DNA or DNA/RNA duplex, the T_m (melting temperature) of the corresponding PNA duplex is increased by 1°C/base or 1.5°C/base, respectively. The further advantage is that mismatches result in greater destabilisation of the duplex which is reflected in a drop in T_m. However more systematic investigation of the chemical and positional effects of mismatches on DNA/PNA duplex stability showed the remarkable stability of some over others (Weiler et al., 1997). By in-gel affinity methods (Igloi, 1998) for an 11-mer PNA showed that destabilisation effect was at maximum if mismatch was at position 4 from either end, but negligible in the case of central position. The most stable mismatched pairs were G:T and T:T whereas maximal destabilisation was for A:A and G:G.

In contrast to DNA, PNA can bind both in parallel and antiparallel orientation, with the classical, 5'- to 3'-end being favoured (Egholm et al., 1993), (Peffer Nancy et al.,). Duplex formation is essentially ionic strength independent which results in the ability of PNA to perform strand invasion (Cherny et al., 1993). These properties of PNA resulted in versatile applications in DNA mutation analysis (Orum et al., 1993), (Demers et al., 1995), affinity capture ((Boffa et al., 1995), and inhibition of gene expression(Good and Nielsen, 1998), (Good and Nielsen, 1998)

All these features make PNA an attractive alternative to DNA oligonucleotide probes in the oligofingerprinting approach. The use of PNA should allow the implementation of probes shorter than 10-mer probes, which form duplexes of similar stability and sufficient discrimination of mismatches. An attractive feature is formation of duplex under low salt condition. At these conditions target DNA is less prone to form secondary structures (e.g. hairpins) and is thus more available for hybridisation. This should result in less "false negative" hybridisation signals. Also since the discrimination of mismatches is alleged to be higher for PNA, the "false positive" rate should be decreased. Finally the use of PNA could result in more accurate clustering and enable the design of motif-specific short oligonucleotide probes to perform gene search based on hybridisation of these.

Although there are no enzymes which use PNA as substrate, it is possible to label PNA radioactively, by chemical modification as in (Kozlov et al., 1998) or else PNAs with modified backbone can by kinased by a protein kinase. Also PNAs are advantageous in multi-step non-radioactive protocols due to their stability. Biotinylated PNA can be produced the same way as oligonucleotides- in the synthesis.

The extreme chemical stability of PNA suggests that wider range of detection methods can be used if compared to conventional oligonucleotides.

1.9 Mass spectrometry as method of hybridisation detection

Mass spectrometry is a powerful method for characterisation of compounds by measuring mass-to-charge ratios of gas-phase ions. The use of MALDI enables to distinguish molecules by their intrinsic physical property of mass-to-charge ratio. This has eliminated the need for both radioactive or non-radioactive labelling and unambiguously defines the compounds. MALDI-TOF (time-of-flight) has many advantages over other mass spec methods. It has lower detection limits, compatible with salts and buffers used in sample preparation and provides

reliable mixture analysis. The most powerful aspect of this technique is high resolution of molecular weights. With mass spectrometry a numerical readout is obtained within seconds with an accuracy of several Dalton. The development of matrix assisted laser desorption-ionisation time-of-flight mass spectrometry has provided an alternative approach to analysis of nucleic acids, however, DNA as a MALDI object has its limitations (Gut et al.,1997).

Compared to DNA, PNA has been described as giving better performance, sensitivity and reproducibility and additionally possess extraordinary stability. PNA used as MALDI analytes have already enabled polymorphism detection based on PNA hybridisation and mass spectrometry detection (Ross et al., 1997), (Griffin et al., 1997). The speed and stringency of PNA hybridisation allow a reduction in time of analysis to minutes. The data resolution can be enhanced by the incorporation mass tags and constructing PNA libraries. It has been suggested (Gut et al, 1998, Figure 1) that combination of PNA and MALDI can be used for DNA clones identification. Coupled with the solid phase affinity assay, MALDI analysis offers speed and high automation and multiplex potential.

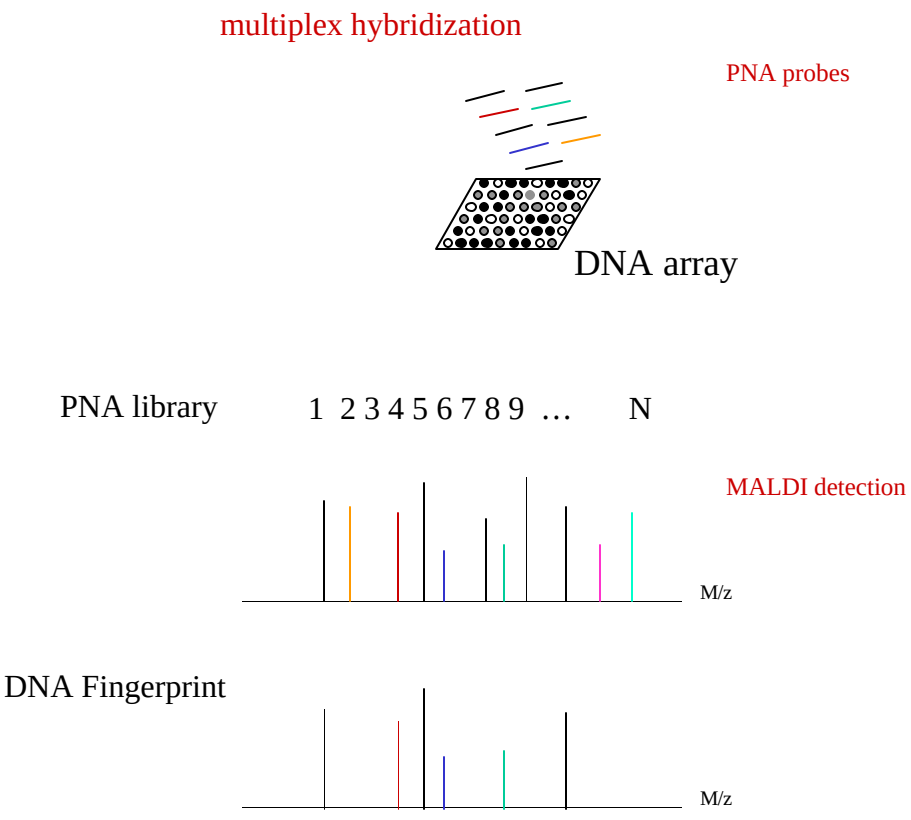


Figure. The concept of OFP with multiplex PNA hybridisation and MALDI detection. A DNA library is hybridised to a PNA library, and a read-out from each dot on an array results in a fingerprint of a correspondent clone.