Chemistry Challenges in SNP Typing

Kazuhiko Nakatani*[a]

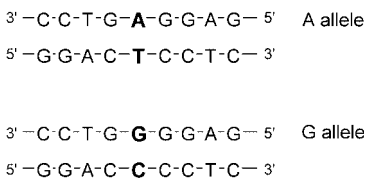
1. Introduction

Single nucleotide polymorphisms (SNPs) are gene variations that result from a single nucleotide difference.[1–4] A single Watson–Crick base pair in wild-type DNA is replaced by any one of the other three base pairs in mutant DNA. When an SNP site is located in a protein-encoding region, the amino acid sequence of the produced protein may have a chance to be changed. SNPs in a promoter region may modulate the transcriptional efficiency of proteins encoded downstream. As a consequence, a small difference in DNA sequence can result in a much larger difference in the phenotype of each individual. SNP typing that identifies the base at predetermined polymorphic sites in any given DNA sample is envisioned as an essential technology in the near feature for realizing personalized medicine. Several methods for SNP typing have been studied, and some of them have already become commercially available. Because there is only a one-base-pair difference between the mutant and wild-type DNAs, it is difficult to differentiate the two DNAs. Most current SNP-typing techniques are indebted to the high fidelity of DNA polymerase in primer extension and flap endonuclease in the digestion reaction of a specific DNA structure. While the difference between the two DNAs is only subtle, increasing knowledge and improving technologies in chemistry and associated fields should lead to a conceptually new solution in SNP typing. Developing nonenzymatic methods for SNP typing offers great opportunities and challenges in chemistry. Among the numerous reports on new methods of SNP typing, this minireview focuses primarily on the chemistry basis of newly developed and developing technologies for SNP typing that do not make use of enzymes, except for PCR amplification. The methods of nonenzymatic SNP typing described here show only the tip of the large number of studies that have appeared in recent scientific journals. Readers who are interested in general methods of SNP typing should consult the comprehensive reviews.[5–9]

The SNP typing strategies described in this review were compared to the principles for allele discrimination, detection signals, and detection platforms (Table 1). While there is no standard condition to compare the sensitivity, accuracy, and performance of each typing method with the others, the amounts of genome, PCR products, or oligonucleotides described for demonstrating the performance of the nonenzymatic methods in the literature are briefly mentioned in Table 1.

2. An Overview of Enzyme-Coupled SNP Typing

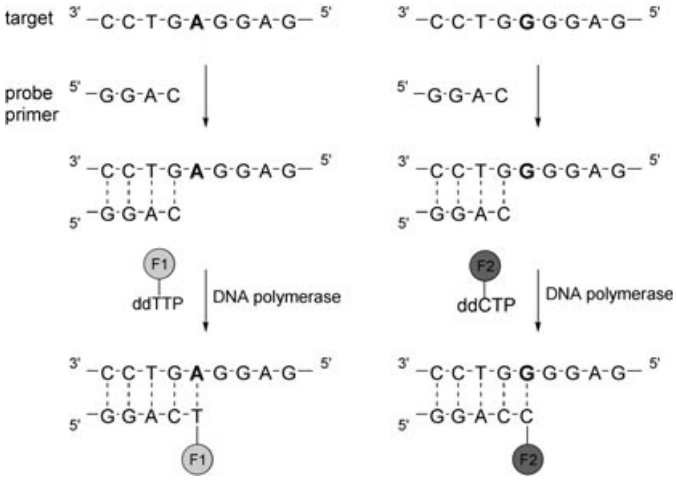
Most current SNP typing methods utilize enzyme reactions as a key transformation in a typing scheme because enzyme-catalyzed reactions show high fidelity in recognition of a single-base difference between wild-type and mutant DNAs. A single primer extension with DNA polymerase[7] and digestion reactions with flap endonuclease[10] are the representative key reactions in enzyme-coupled SNP typing. Scheme 1 shows a typical example of SNPs. The wild-type has an A–T base pair at the polymorphic site, whereas the mutant DNA has a G–C base pair.



Scheme 1. *Single-nucleotide polymorphisms. Wild-type DNA (upper) has an A–T base pair, whereas mutant DNA has a G–C base pair at the polymorphic site.*

2.1. Single-nucleotide primer extension

A general scheme for SNP typing with a single nucleotide primer extension is shown in Scheme 2.[7]



Scheme 2. *Single-nucleotide primer extension of wild-type and mutant DNAs with labeled ddNTPs.*

[a] *Dr. K. Nakatani*
Department of Synthetic Chemistry and Biological Chemistry
Graduate School of Engineering, Kyoto University
Kyoto 615-8150 (Japan)
Fax: (+81) 75-383-2759
E-mail: nakatani@sbchem.kyoto-u.ac.jp

 DOI: 10.1002/cbic.200400161

Table 1. *Comparison of SNP typing strategies.*

Typing Methods	Principle for allele discrimination	Detection signals	Detection platform	Advantages and sensitivity	Refs.
Enzyme-coupled methods					
single primer extension	enzymatic base-sequence recognition in primer extension	fluorescence and mass spectrometry	array and homogeneous	suitable for detection by MALDI-TOF	[7, 8, 11–13]
Invader assay®	sequence specific cleavage by Flap endonuclease	fluorescence	homogeneous	commercially available	[10]
TaqMan assay®	allele-specific hybridization	fluorescence	homogeneous	commercially available	[14, 16]
molecular inversion probe	allele-specific gap-filling reactions	fluorescence	array	typing both alleles with single MIP probe. Use four enzymes.	[19]
Nonenzymatic methods					
strand exchange	hybridization kinetics accelerated by cation comb-type copolymer	fluorescence	homogeneous	a unique discrimination principle. 80-mer single-stranded target (0.6 pmol)	[22]
redox-active intercalator	allele-specific hybridization with a combination of duplex-selective drug binding	differential pulse voltammetry	array	commercially available, 0.2 pmol of PCR products	[24, 25]
charge transport	low efficiency in charge transport through base mismatches	chronocoulometry	array	signal enhancement by catalytic redox cycles. ~10^8 duplexes could be detected.	[27–29]
Magi Probes®	modulated intercalative binding of quencher by the neighboring base pairs	fluorescence	homogeneous array	30 pg of genomic DNA was used in real-time PCR assay	[30]
modified nucleotide bases	modulation of microenvironment of fluorophore upon hybridization, base discrimination by hydrogen bonding	fluorescence	homogeneous array	oligomer duplex (~2 μM) was used for homogenous assay	[34–36]
mismatch binding ligands (MBL)	recognition of base mismatches in heteroduplexes by specific binding of ligands	surface plasmon resonance	array and micro channel	free from oligonucleotide labeling, minimum sample (27-mer duplex) was 1~10 nM in100 μL (0.1~1 pmol)	[37, 40–43]

The probe primer that is common for both the wild and the mutant targets hybridizes on the 3′ side to a polymorphic site. For the fluorescence detection method, dideoxynucleotide triphosphates (ddNTPs) that are to be incorporated into a primer are labeled with fluorescent dyes that have different emission maxima. In the primer extension, ddTTP is incorporated into the 3′ end of the primer hybridized to the wild-type, whereas ddCTP will be incorporated into the primer hybridized to the mutant. Observation of fluorescence specific to ddTTP or ddCTP in the extended primer defines the base at the polymorphic site. Alternatively, the extended products can be directly analyzed by matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF-MS).[8, 11–13] MALDI-TOF-MS is capable of ionizing DNA without producing fragments, and can determine the molecular weight of the elongated primer. Because the mass difference between two primers into which T and C are incorporated at the 3′ end is only 15 Da, separation of the two products needs high resolution. Chemically modified ddNTPs with molecular tags are used to increase the mass difference between the products and, hence, to gain unambiguous typing results. In this respect, ddCTP and ddTTP labeled with different fluorescent dyes could also be used for MALDI-TOF-MS detection.

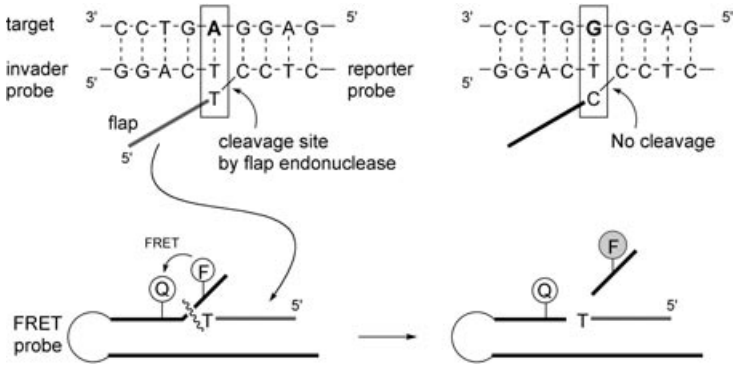


Figure 1. *Schematic representation of the Invader® assay.*

2.2. Invader® assay

One method that effectively uses the enzyme reactions in SNP typing is Invader® assay (Figure 1).[10] The assay utilizes two oligonucleotides, called "reporter" and "invader" probes, and an enzyme called flap endonuclease. The reporter probe consists of an allele-specific sequence and an additional sequence called a flap at the 5′ side of the probe. The reporter probe hybridizes to the target on the 5′ side to the SNP site at the allele-specific sequence. The flap sequence is not related to the target sequence at all. The invader probe is also allele specific and hybridizes to the 3′ side to the SNP site. The nucleotide at the end of the invader probe is located just opposite the nucleotide to be identified in the target. When the invader and the reporter probes, which are both specific to the wild-type, hybridize to the wild-type DNA, two thymine bases are located opposite the adenine at the SNP site. In contrast, a cytosine base in the reporter probe and a thymine base in the wild-type specific invader probe are located opposite a gua-

nine at the SNP site in the mutant DNA. The flap endonuclease recognizes the former DNA structure and cleaves the reporter probe at the base opposite the SNP site. In Invader assays, the cleaved flap sequence is designed to be another invader probe in a signal-producing system. The flap sequence in the second reporter probe (FRET probe) is labeled with a fluorescent dye. The fluorescence of the dye is quenched by fluorescence resonance energy transfer (FRET) to the quencher attached to the stem part in the reporter probe. Cleavage of the flap in the FRET probe by the flap endonuclease results in the disruption of FRET, and thus an increase in the fluorescence. As a hybridized complex produced from the FRET probe specific to the wild type is not recognized by the enzyme, the intensities of the fluorescence signals do not increase. Cleavage of the flap is a catalytic process, and the cleaved probe can be displaced by other reporter probes. In principle, it is conceivable that Invader assays could be carried out without PCR amplification of the target prior to the assay.

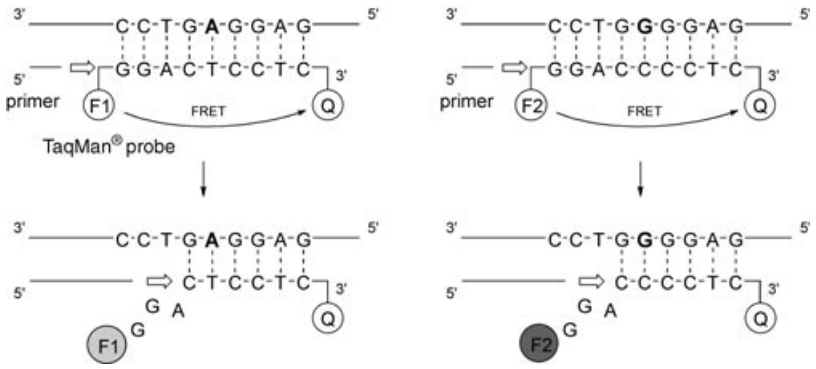


Figure 2. *The TaqMan® assay.*

2.3. TaqMan® assay

Another important method of enzyme-coupled SNP typing is TaqMan® assay (Figure 2).[14] This assay is a kind of allele-specific hybridization method[15] coupled with PCR amplification. In addition to a primer set for the PCR reaction, this assay uses TaqMan probes that are allele specific and labeled with fluorescent dyes and quenchers. The TaqMan probe hybridizes to the sequence containing the polymorphic site in the target. In the hybrid, the fluorescence of the dye is quenched by FRET. A primer hybridizes to the 3′ side to the polymorphic site and is elongated by the Taq polymerase. During the polymerase reactions, the TaqMan probe is replaced by an elongating primer and digested by exonuclease activity of the polymerase; this results in the termination of FRET by separation of the fluorescent dye from its quencher. The TaqMan probe specific to the wild-type DNA does not hybridize to the mutant DNA, especially in the presence of the competitive mutant specific probe, and, therefore, will not be digested by the enzyme in the absence of the consensus target DNA. Because the first step of this assay is the allele-specific hybridization (discussed in detail in Section 3), discrimination between wild-type and mutant targets by the probe was done by thermodynamic difference between a fully matched duplex from a single mismatch duplex. In the case of A–T-rich target sequences, the probe length should be long enough to gain thermal stability of the probe–target duplex; this makes the FRET efficiency decrease due to a long distance between fluorophore and quencher. In order to circumvent these difficulties by increasing the difference in the thermodynamic stability between matched and mismatched duplexes without increasing the probe length, an improved method utilizing an MGB (minor-groove binder) probe has been developed.[16]

2.4. Ligase-assisted assay

Ligase-assisted assay is one of the important methods for enzyme-coupled SNP typing.[17,18] A unique typing method that involves circular DNA formation by a combination of gap-fill polymerization and a subsequent ligation has been developed.[19] The assay uses molecular inversion probes (MIPs; Figure 3). The MIP consists of two regions hybridized to the

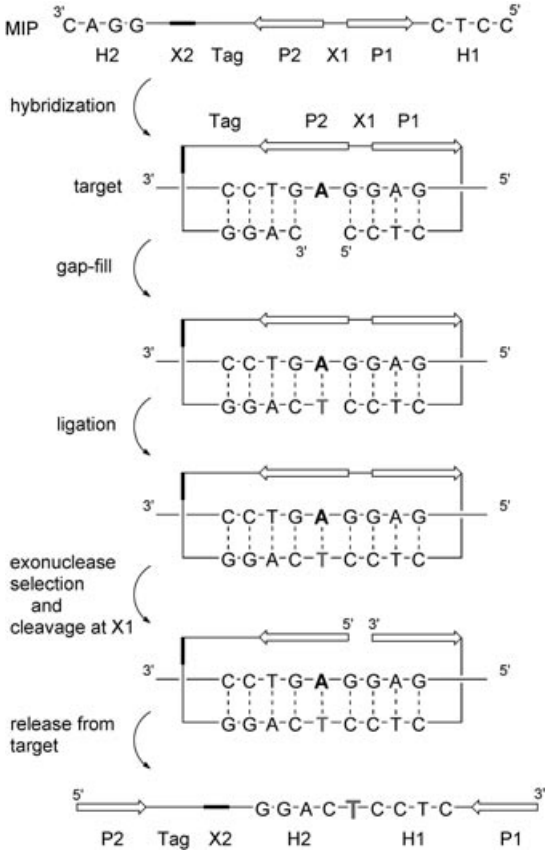


Figure 3. *The molecular inversion probe assay.*

target sequence (H1 and H2), two primer sites (P1 and P2), and a Tag sequence hybridized to the arrayed sequence in the detection step. The sites and X1 and X2 are incorporated to cleave the circularized probe. X1 in the Figure denotes a uracil, which is removed by uracil DNA glycosylase. Upon hybridization of MIP to the target, a gap appears between H1 and H2 and located opposite the polymorphic site. The gap is filled by polymerization with each one of four dNTPs followed by ligation to produce a circular probe. The gap is not filled with uncomplimentary dNTPs to the nucleotide base at the polymorphic site. The remaining probes in a linear form are digested by exonuclease treatments. Having completed the selection, the circularized probe is cleaved at the X1 site (uracil) by uracil DNA glycosylase; this makes the probe inverted to the linear form, with two primer sites being located at each termini. The inverted probe is amplified by PCR and hybridized to the arrayed sequence at the Tag sequence. Hybridization is detected by use of a fluorescence-labeled oligonucleotide complementary to one of primer sequence. This assay needs only one probe for typing two alleles, thus making multiplex genotyping extremely simple.

3. Issues in Nonenzymatic SNP Typing

The major advantage of enzyme-coupled SNP typing is the high fidelity of the base sequence recognition. This implies that nonenzymatic SNP typing methods should be comparable in their ability to recognize base sequences or should use other principles for the efficient discrimination of the mutant DNA from the wild-type. Allele-specific hybridization[15] is a standard method employed in nonenzymatic discrimination of the mutant from the wild-type. This method utilizes allele-specific oligonucleotides (ASO) that hybridize to the target sequences containing the polymorphic site. Hybridization of ASO that is specific to the wild-type target produces a fully matched duplex, whereas a duplex containing a single mismatched site would be produced upon hybridization with mutant DNA. The two duplexes exhibit a different thermodynamic stability due to the presence of a single mismatched site, showing a lower melting temperature (T_m) for the mismatched duplex than the T_m of a fully matched duplex. Under appropriate conditions, a mismatched duplex can be denatured to single strands whilst keeping the most part of the fully matched duplex in a duplex form. With a fluorescently labeled target and ASO immobilized to the solid surface, a labeled target in a mismatched duplex could be washed away from the surface and a fully matched duplex can be easily detected.

One critical issue in the ASO hybridization method is the difficulty in choosing the appropriate conditions for hybridization that discriminate a fully matched duplex from a single mismatched duplex. To circumvent the issues, a new SNP scoring method, dynamic allele-specific hybridization (DASH) has been developed.[20] In simultaneous and multiple SNP-typing formats, the hybridization conditions are especially difficult to set up for the many different ASOs of varied sequences. Careful selection of the sequence and G–C contents to make the thermodynamic stability of the producing duplexes consistent is essential for the design of large multiplex assay formats. Thus, discrimination of matched and mismatched duplexes by their thermodynamic stabilities is not likely to be a useful chemical basis for nonenzymatic SNP typing.

Another important issue regarding nonenzymatic SNP typing is cost. Most typing methods utilize fluorescence detection, which requires fluorescent labeling of either the target DNA or hybridization probes. Fluorescent labeling of target DNA can be easily carried out during PCR amplification with fluorescently labeled dNTPs or primers. The latter could be synthesized by standard oligo DNA synthesis with fluorescently labeled phosphoramidites. However, fluorescently labeled dNTPs and phosphoramidites are expensive, and potentially increase the total costs of SNP typing. Chemically modified DNA oligomers may be used as a substitute for ASO in nonenzymatic SNP typing,[21] but the cost of custom syntheses of these modified probes and their quality control would be potentially an impediment for general use. The use of unlabeled oligomers is one of the ultimate chemistry challenges.

The last issue in nonenzymatic SNP typing is the necessity for PCR amplification prior to the assay. As discussed above, Invader assay could be carried out without PCR amplification of the target DNA prior to the assay because digestion of the specific DNA complex is enzyme catalyzed. Amplification of the target DNA is not possible without the enzyme. Under low concentrations of the target DNA, the signal intensity remains weak. One possible solution to this issue is to amplify the detecting signal instead of amplifying the sample DNA. In this respect, electrochemical detection is an attractive approach because of the high sensitivity and possible signal amplification.

It is important to note that two oligonucleotide primers are needed as long as PCR is used prior to the assay. Thus, ASO and other methods using any type of allele-specific hybridization probes require at least two additional oligonucleotides.

4. Nonenzymatic SNP Typing Methods

4.1. Strand-exchange methods

As discussed in a previous section, it is difficult in practice to define the hybridization conditions for SNP typing by ASO hybridization,[20] especially for large-scale and simultaneous assay formats. A unique SNP-typing method employing a kinetic rather than a thermodynamic differentiation of a fully matched duplex from a single mismatched duplex has been reported (Figure 4).[22] The method is based on the discovery that a cation comb-type copolymer (CCTC) accelerates the strand exchange of a fully matched duplex with a matched complementary strand but not with a single mismatched strand. The method utilized a double-stranded probe, in which one strand was labeled with FITC (fluorescein isothiocyanate) and the other complementary strand was labeled with carboxytetramethylrhodamine (TAMRA). In the double-stranded probe, the fluorescence of FITC was suppressed by FRET to TAMRA. The double-stranded probe was incubated with the target sample that can hybridize to the FITC-labeled strand by displacing the TAMRA-labeled strand. CCTC accelerates the strand-exchange

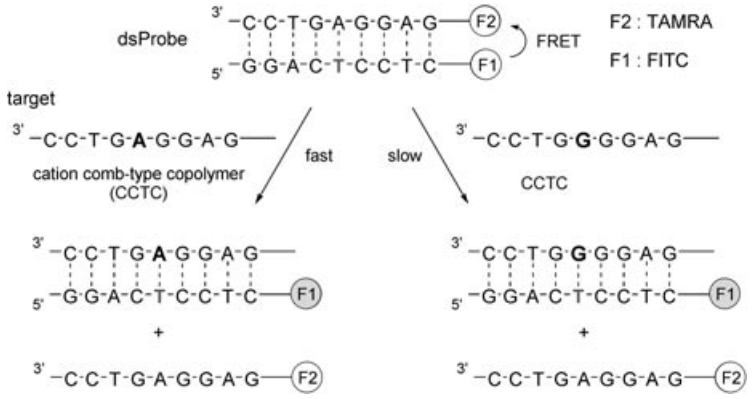


Figure 4. *Strand-exchange methods with unlabeled targets.*

reaction when the target has the sequence that is fully complementary to the FITC-labeled strand. Strand exchange resulted in an increase in the fluorescence intensity of FITC by termination of FRET. The time-course measurements of increasing intensity determine the allele type.

4.2. Electrochemical detection

Due to the expected high sensitivity and capability for signal amplification, electrochemical detection of SNP is an attractive option and has been intensively studied. An excellent review has appeared recently.[9] The strategies for SNP typing by electrochemical detection have been classified into five categories: 1) direct and 2) indirect DNA chemistry, 3) DNA-specific redox indicator detection, 4) nanoparticle-based electrochemistry amplification, and 5) DNA-mediated charge transport. The advantages and disadvantages of each strategy were discussed in detail in the review.

As the purine bases can be electrochemically oxidized, the amount of free purines determined by adsorption-stripping voltammetry can be used as an index of the amount of the captured target DNA by ASO hybridization. The indirect method utilizes electrochemical mediators for the oxidation of guanine bases. Polypyridyl complexes of Ru^{II} and Os^{II} have been successfully used for this method.[23] Since these methods are based on ASO hybridization, difficulties associated with ASO hybridization methods are inherently involved. As an analogy with fluorescent labeling, chemical labeling of ASO with electrochemically active reporter molecules has also been studied for SNP typing.

4.2.1. Methods using DNA-binding ligands with redox-active molecules: Instead of direct labeling of ASO, an alternative method that makes use of a DNA intercalator containing redox-active reporter molecules has been studied (Figure 5).[24, 25] Takenaka et al. have synthesized an intercalator naphthalene diimide tethered with two ferrocenyl groups. Upon intercalation of naphthalene diimide into the DNA duplex, each pair of bulky ferrocenyl groups is located one in a minor and the other in a major groove of DNA; thus dissociation of the intercalators from the duplex is 100 times slower than from single-stranded DNA. This results in an increase in binding selectivity of this intercalator to the duplex. As binding of the ferrocenylnaphthalene diimide to single-stranded DNA increases the background current level, duplex selectivity for the intercalator is essential to discriminate the duplex from single-stranded DNA. In this regard, the amount of the mismatched duplexes should be kept to a minimum by choosing appropriate washing conditions, because the intercalator potentially stabilizes both the mismatched and matched duplexes. Hybridization of ASO immobilized on a gold electrode to the target DNA can be monitored by current responses from the ferrocenylnaphthalene diimide intercalated into the duplex.

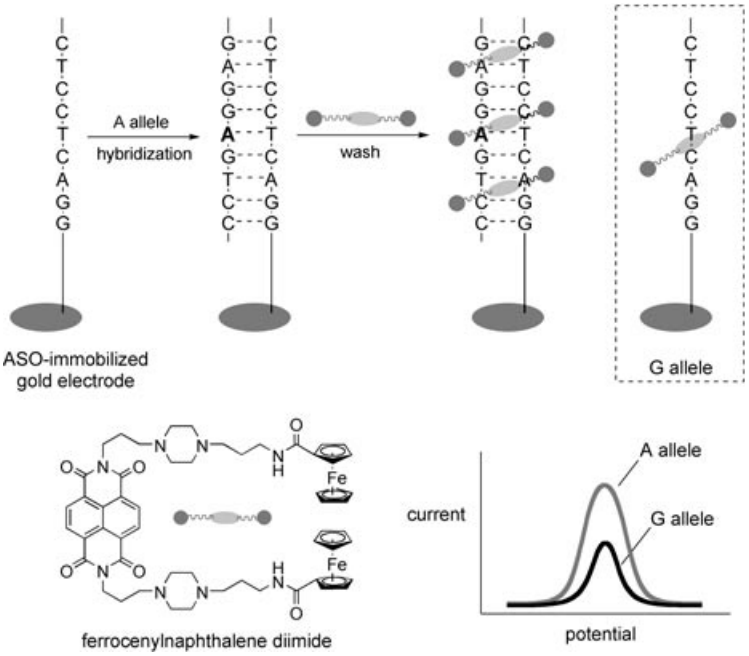


Figure 5. *Electrochemical detection with ferrocenylnaphthalene diimide.*

4.2.2. Methods based on charge transport through the DNA π-stacking: Electrochemical labeling of ASO and intercalators has been used to determine the amount of DNA and whether it is double-stranded or single-stranded. On the basis of DNA-mediated charge-transport studies,[26] a unique method sensing the structural perturbations of the π-stacking at the SNP site has been developed.[27] It has been demonstrated that disruption of the π-stacking of base pairs significantly reduces the charge-transport efficiency through DNA.[28] Charge transport was selectively terminated at the mismatched base pairs, where the duplex π-stacking was structurally perturbed. Barton and co-workers utilized a densely packed duplex monolayer immobilized on a gold electrode, where the DNA duplex was expected to orient itself in an upright position on the

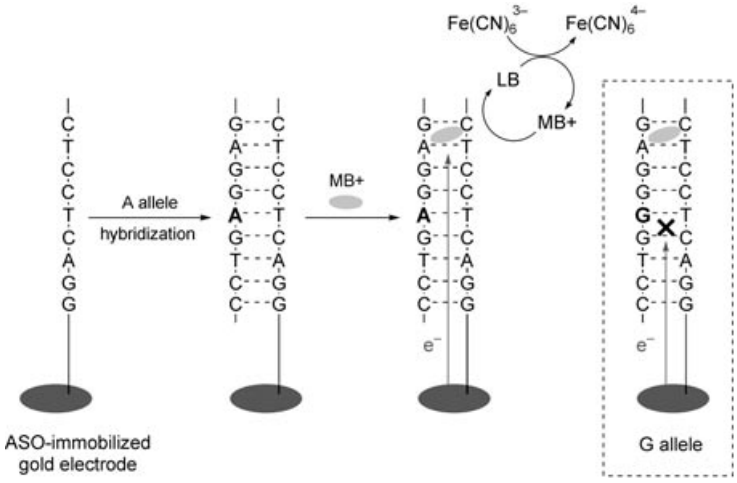
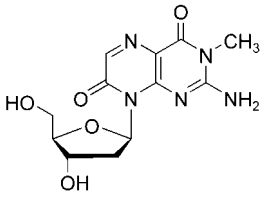


Figure 6. *Electrochemical detection by charge transport through duplex π-stacking.*

gold surface due to the high DNA density (Figure 6). For the monolayer, the binding of DNA intercalators was primarily restricted to the top of the surface. They used methylene blue (MB+) as a redox-active DNA intercalator and coupled the electrochemical reduction of methylene blue with the electrocatalytic reduction of ferricyanide by the reduced form of methylene blue (leucomethylene blue, LB). In the electrocatalytic cycle, electrons flow from the gold electrode to methylene blue through the duplex π-stacking to produce leucomethylene blue. Ferricyanide ions in the solution were then reduced by leucomethylene blue at the surface to regenerate methylene blue. When a mismatched duplex is produced on the gold electrode, the electron flow from the electrode is disturbed at the mismatched site due to improper stacking of the mismatched base pair to the neighboring base pairs. It was reported that the presence of the mismatch caused a sixfold decrease in the electrocatalytic current.[29]

ing dynamics of DNA.[31] Because the fluorescence is sensitive to the microenvironment surrounding the fluorescence chromophore, these nucleotide base analogues have been utilized to report the hybridization. 3-Methyl isoxanthopterin (3-MI) has been synthesized and incorporated into oligonucleotides (Scheme 3).[32] With a combination of bulge-forming hybridization, the fluorescence of 3-MI incorporated into oligonucleotides showed a 27-fold increase compared to the single-stranded state. It is reported that the fluorophore bulged out from the double strand upon hybridization; this result-



Scheme 3. *Fluorescent pteridine nucleoside.*

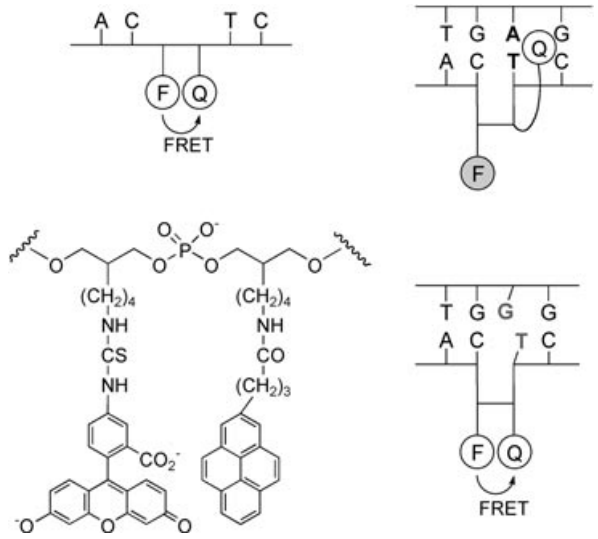


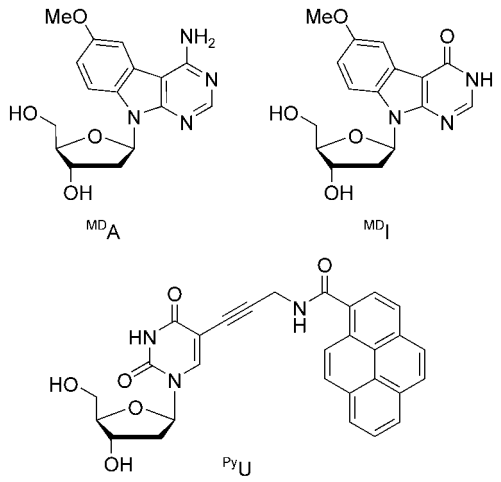
Figure 7. *FRET detection by modified ASO.*

4.3. Chemically modified fluorescence-probe methods

4.3.1. Based on FRET chemistry: Oligomers that have a fluorescence probe and intercalating quencher in close vicinity on the same strand have been developed for discriminating a perfectly matched duplex from a singly mismatched duplex (Figure 7).[30] In a single-stranded state, the fluorescence of fluorescein was quenched by a quencher pyrene. When hybridized to the perfectly matched sequence, pyrene intercalated into the duplex; this resulted in the emission of fluorescence from fluorescein due to the disruption of FRET. In the case of hybridization with a mismatched sequence, pyrene intercalation is inhibited by the mismatched base pair. As a consequence, the fluorescence remains quenched. These probes have been termed MagiProbe®. The performance for match–mismatch discrimination is modulated by the ability of the intercalative binding of the quencher.

4.3.2. Modified nucleotide-base methods: Fluorescent nucleotide-base analogues have been investigated as probes reported in increased fluorescence because the quenching of the fluorescence by the neighboring bases was no longer effective.[33]

A series of novel modified nucleotide bases that emit fluorescence when they hybridize to a specific nucleotide base has been reported (Scheme 4).[34] Methoxybenzodeazaadenine (^{MD}A) emits fluorescence at 397 and 427 nm upon excitation at 330 nm with a quantum yield of 0.118. When ^{MD}A was incorporated into single-stranded DNA oligomers, the fluorescence was very weak. In a fully matched duplex in which thymine was located opposite ^{MD}A, the fluorescence was also weak. However, the fluorescence became 100 times stronger when cytosine was incorporated opposite ^{MD}A than was observed in fully matched duplexes. Methoxybenzodeazainosine (^{MD}I) has similar fluorescent properties. It emits fluorescence at 424 nm in a duplex in which thymine is located opposite ^{MD}I. With the ASO incorporating these modified bases, SNP typing of the human breast cancer 1 gene has been demonstrated. The dif-



Scheme 4. *Fluorescence nucleotide bases.*

ference between this method and conventional SNP typing with ASO hybridization is that thermodynamic discrimination of matched and mismatched duplexes is unnecessary. As discussed above, the fluorescent labeling of ASO has the impediments of the cost of synthesis and the need for many kinds of labeling oligomers in practice. In addition, the quenching of fluorescence of these modified bases by the neighboring guanine limits the applicable sequence. While solutions for these issues are essential, the method is potentially suitable for the homogeneous-assay format of multiple high-throughput SNP typing. Different types of artificial nucleosides containing pyrenecarboximide (^{Py}U), which allow fluorescent quenching by the neighboring guanines to be circumvented, have been reported.[35] The ^{Py}U–A base pair emitted strong fluorescence at 397 nm upon excitation at 327 nm, whereas the fluorescence from ^{Py}U–N (N = C, G, or T) was considerably weaker. The fluorescence of the ^{Py}U–A base pair was not quenched by a flanking G–C base pair. The group also developed pyrenecarboximide-labeled cytidine.[35]

The fluorene-modified uridine (Scheme 5) incorporated into the loop region of the hairpin duplex has been used for SNP typing.[36] Upon hybridization with a perfectly matched target

Scheme 5. *Fluorene-modified uridine.*

to produce a ^{Fl}U–A base pair, the intensity of fluorescence at 425 nm by excitation at 340 nm showed a 2.2-fold increase. In contrast, the formation of a ^{Fl}U–C mismatched base pair resulted in a 0.15-fold decrease in the fluorescence intensity. The authors claimed the discrimination factor to be 14.7. It is noteworthy that the fluorescence of ^{Fl}U was not quenched by the flanking G–C base pairs.

4.4. Methods based on mismatch-binding ligands

4.4.1. Heteroduplex analysis: We[37–43] and others[44–46] have pursued a completely different approach to chemical SNP typing from those utilizing chemically modified ASO or ASO-immobilized surfaces as discussed above. One of the challenges we have focused on is the reduction and eventual redundancy of labeled oligonucleotides for analysis. The expense of the fluorescent labeling of oligonucleotides and PCR products, the synthesis of chemically modified ASO, and the covalent immobilization of ASO to the surfaces potentially limit their applications in SNP typing. The principle of SNP typing we have focused on is heteroduplex analysis,[47] which is well recognized as a method for mutation detection. We have developed an integrated heteroduplex analysis by combination with mismatch-binding ligands (MBL). Standard wild-type DNA and test DNA extracted from patients' blood were mixed and simultaneously amplified by PCR with one primer set. Note that the two DNAs can be amplified with the same primer set, because wild-type and mutant DNA have the same primer sequence. After PCR amplification, the produced duplexes were denatured and then annealed. As the duplexes produced from test and standard DNA samples differed from each other by a single nucleotide sequence, hybridization of two sets of duplexes produces DNA heteroduplexes containing a single mismatched site in addition to the completely matched homoduplexes (Figure 8). Mismatch-containing duplexes can be separated from homoduplexes either by gel electrophoresis and chemical and enzymatic cleavages at the mismatched site or by selective capture with mismatch-binding proteins. While these heteroduplex analy-

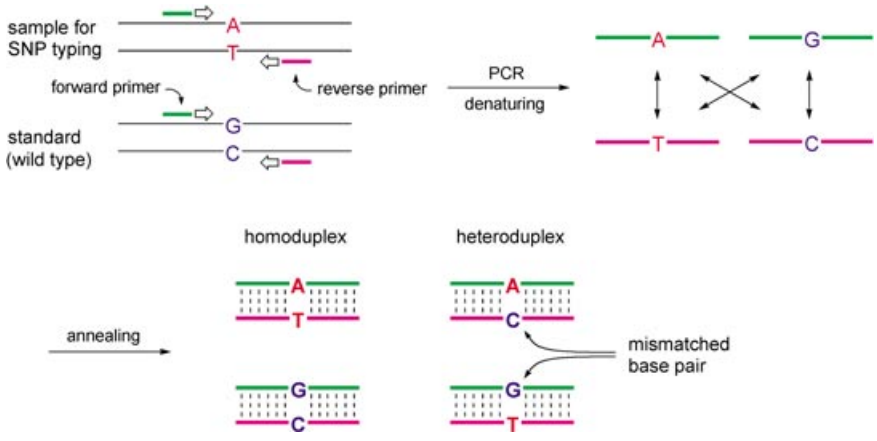
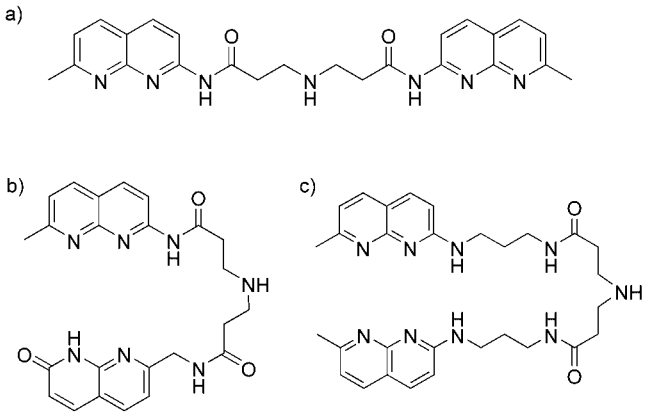


Figure 8. *Schematic representation of heteroduplex analysis.*

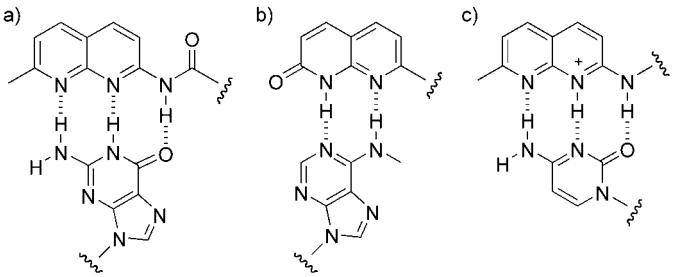
ses as applied to low-throughput screening are essentially free from oligonucleotide labeling, new technologies for high-throughput analyses are yet to be established. Mismatch-binding ligands that selectively bind to a mismatched site in heteroduplexes could replace mismatch-binding proteins and bring an innovation to heteroduplex analyses.

4.4.2. Mismatch-binding ligands: We have so far succeeded in developing three MBLs (Scheme 6). The naphthyridine dimer (**ND**),[37,38] which consists of two 2-amino-7-methyl-1,8-naphthyridines and a linker connecting the two chromophores, strongly and selectively binds to guanine–guanine mismatches. The dissociation constant (K_d) for the 1:1 complex of **ND** and



Scheme 6. *Mismatch binding ligands (**MBL**) a) naphthyridine dimer (**ND**), b) naphthyridine-8-azaquinolone (**NA**), c) aminonaphthyridine dimer (**amND**).*

the G–G mismatch in the 5′-CGG-3′/5′-CGG-3′ sequence was about 100 nM by isothermal titration calorimetry. Naphthyridine azaquinolone (**NA**),[40] in which one naphthyridine chromophore in **ND** was replaced by 8-azaquinolone, showed affinity selective to the guanine–adenine mismatch. The affinity and stoichiometry for **NA** binding to G–A mismatches depend on the sequence flanking the mismatch. The aminonaphthyridine dimer (**amND**),[41] in which two naphthyridine chromophores are tethered by an alkyl linkage, showed preferential binding to mismatches that consisted of two pyrimidines, especially cytosine–cytosine mismatches. Each of the three chromophores in these MBLs has a hydrogen-bonding surface complementary to that of a target nucleotide base in the mismatched base pairs. Upon binding of these MBLs to the target mismatch, the chromophores produce a hydrogen-bonded pair with the mismatched bases (Scheme 7). The naphthyridine chromophore, having an acceptor–acceptor–donor alignment of hydrogen-bonding groups, is fully complementary to that of guanine, whereas a surface with an acceptor–donor–acceptor alignment, as in 8-azaquinolone, is complementary to that of adenine. The aminonaphthyridine chromophore is not complementary to cytosine in the hydrogen-bonding surface, but is thought to be protonated at pH 7, thus producing the complementary hydrogen-bonding surface to that of cytosine.[41] The resulting pseudo base pairs between MBL and the mismatched base pairs would be stacked by the flanking base pairs in a duplex π-stack, resulting in formation of stable MBL–DNA complexes.



Scheme 7. *Possible hydrogen bonding between a) naphthyridine and guanine, b) 8-azaquinolone and adenine, and c) protonated naphthyridine and cytosine.*

4.4.3. Surface plasmon resonance assay with an MBL-immobilized sensor: We have developed novel sensors for a surface plasmon resonance (SPR) assay to detect the mismatch duplexes. The mismatch-detecting sensors were prepared by immobilizing MBLs on the surface of commercially available SPR sensor chips (Figure 9). SPR detects the change in the refractive index caused by variation of the mass on the sensor-chip surface, for example, when the analyte binds to the immobilized ligand on the surface.[48] The change in SPR signal, termed the SPR response presented in resonance units (RU), is directly related to the change in surface concentration of biomolecules. An SPR response of 1000 RU is equivalent to a change in surface concentration of 1 ng mm^{-2}. Thus, the density of immobilized ligands on the surface and amount of analyte bound to the surface can be calculated from the difference in SPR response before and after the analyses. The SPR signal is monitored continuously, so that chemical interactions between ligand and DNA can be studied in real time. MBLs were attached to an adapter linker at the secondary amino group by an alkyl linkage. Acylation of the amino group increased the rigidity of the linker and resulted in a decrease in affinity for the mismatch. The resulting tertiary amino group is protonated at a neutral

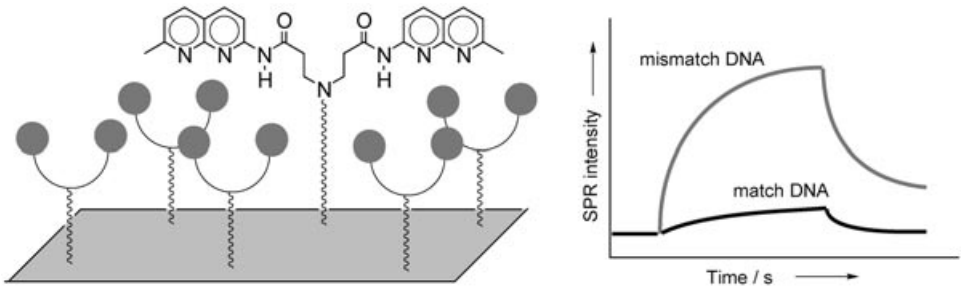


Figure 9. *An image of the MBL-immobilized SPR sensor and simulated SPR responses to DNA.*

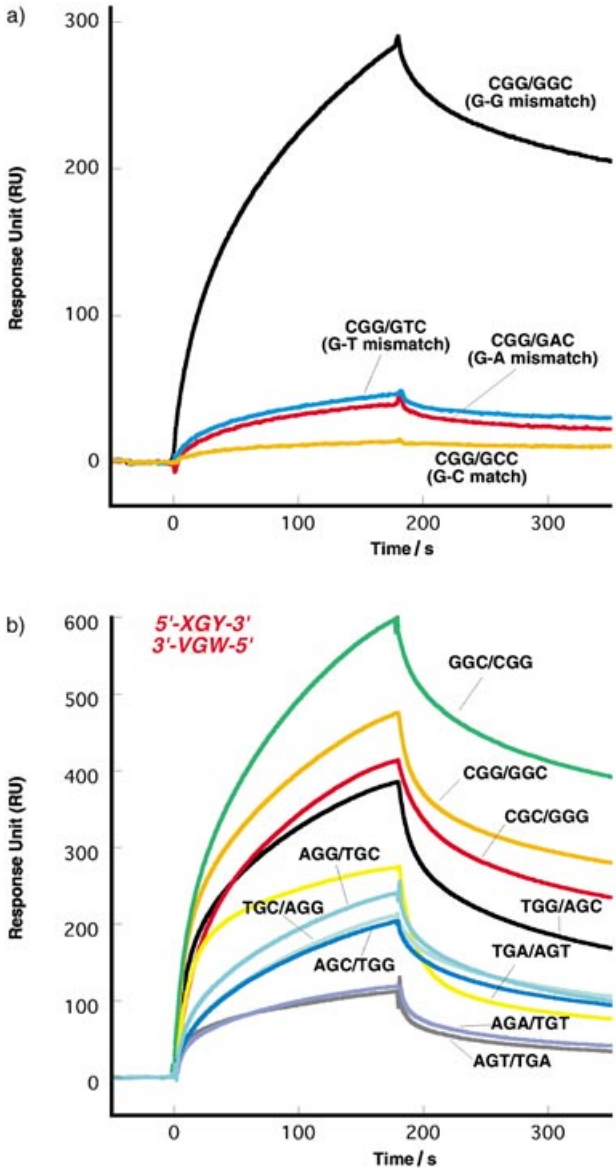


Figure 10. *SPR analyses of a) G–G, G–A, and G–T mismatches containing 27-mer duplexes and b) G–G mismatch containing duplexes with* ***ND****-immobilized sensor.*

pH and, hence, increases the attractive electrostatic interactions between the ligand and DNA. The terminal amino group of the adapter linker could be attached to the carboxyl group tethered to the SPR surface either directly or by way of a poly(ethylene oxide) (PEO) linker. Increased signal intensities in SPR analyses were observed by incorporating PEO linker units.[42]

The **ND**-immobilized sensor surface produces an intense SPR response selective to the duplex containing a single G–G mismatch (Figure 10a).[37] Weak SPR responses were obtained for G–A and G–T mismatches, whereas the fully matched duplex showed an insignificant response under the same conditions. The observed mismatch selectivity is quite consistent with the results obtained by the thermodynamic stabilization of the mismatch by **ND**. 27-mer duplexes 5′-d(GTTACAGAATCT**W**AAGCCTAATACG)-3′/3′-d(CAATGTCTTAGA**XGZ**TTCGGATTATGC)-5′ containing G–G mismatches with all possible flanking sequences were analyzed by an **ND**-immobilized sensor surface (Figure 10b). The results showed that binding of **ND** to G–G mismatches was affected by the base pairs flanking the mismatch. G–G mismatches flanking G–C base pairs showed stronger SPR responses than those flanking A–T base pairs. This is most likely due to a stabilization of the pseudo base pairs by stacking with the flanking G–C base pairs. As the fully matched duplex showed only an insignificant response, G–G mismatches could be detected with an **ND**-immobilized sensor surface irrespective of the flanking sequence. With incorporation of three PEO units in the linker, we could successfully detect a G–G mismatch in the 27-mer duplex at a concentration of 1 nM.

The **NA**-immobilized sensor showed affinity for the G–A mismatch and also for the G–G mismatch with lower efficiency (Figure 11a).[40] This is most likely due to a possible tautomeric 2-hydroxy-1,8-naphthyridine form of 8-azaquinolone, which has a hydrogen-bonding surface partially complementary to that of guanine. In contrast to the strong **ND** binding for all G–G mismatches, **NA** binding to G–A mismatches was found to be highly sequence dependent. SPR analyses of 27-mer duplexes containing a G–A mismatch with the **NA**-immobilized sensor surface showed that G–A mismatches in the sequence of 5′-TGG-3′/3′-AAC-5′, CGG/GAC, TGC/AAG, GGA/CAT, AGG/TAC, and AGA/TAT could be detected, but it was not possible to differentiate G–A mismatches in other sequences from the fully matched duplex (Figure 11b). Highly sequence-dependent binding of **NA** to the G–A mismatches implies that **NA** binding involves interactions not only to the G–A mismatch but also to the base pairs flanking the mismatch. Because sequence-dependent binding of MBLs to the mismatch is a drawback with respect to SNP typing, an improved molecular design is necessary for the MBLs that strongly bind to the G–A mismatches.

The **amND**-immobilized surface showed a strikingly different binding property from **ND**- and **NA**-immobilized surfaces with respect to the detectable mismatches (Figure 12a).[41] A C–C mismatch showed an intense SPR intensity, whereas the fully matched duplex failed to produce any significant SPR. In addition to the C–C mismatch, a C–T mismatch also produced a significant SPR signal with reduced efficiency. Substitution of an acyl linkage in **ND** with an alkyl linkage in **amND** dramatically changed the preference of the MBL from purine to pyrimidine mismatches. In **amND**, the naphthyridine chromophore was electron rich and susceptible to protonation at neutral pH. When the C–C mismatch duplex was analyzed with the **amND**-immobilized surface at pH 8, SPR signals were found to be very weak compared with the signal at pH 7. SPR analyses of 27-mer duplexes containing C–C mismatches with an **amND**-immobilized sensor surface showed a sequence dependence for the **amND** binding (Figure 12b). Again, C–C mismatches flanking the G-C base pairs produced a stronger SPR signal than those flanking A–T base pairs. As the background signal due to a fully matched duplex was extremely weak, all C–C mismatches irrespective of flanking base pairs could be detected by the **amND**-immobilized surface.

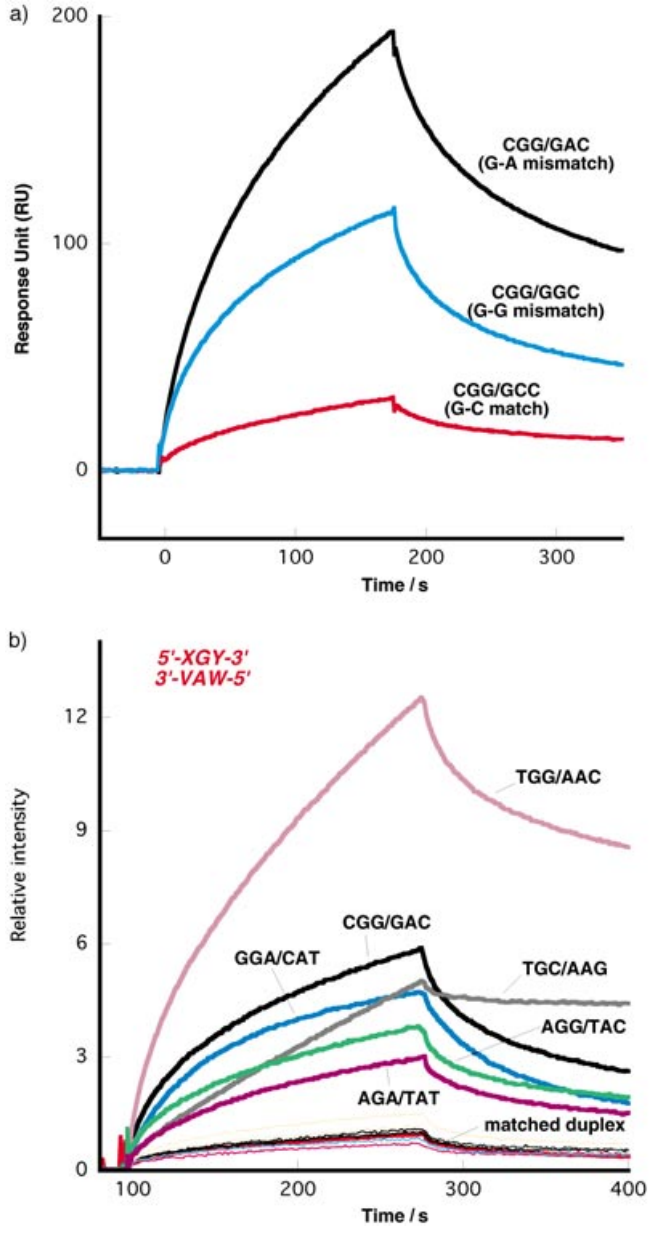


Figure 11. *SPR analyses of a) G–A and G–G mismatches containing 27-mer duplexes and b) G–A mismatch containing duplexes with* ***NA****-immobilized sensor.*

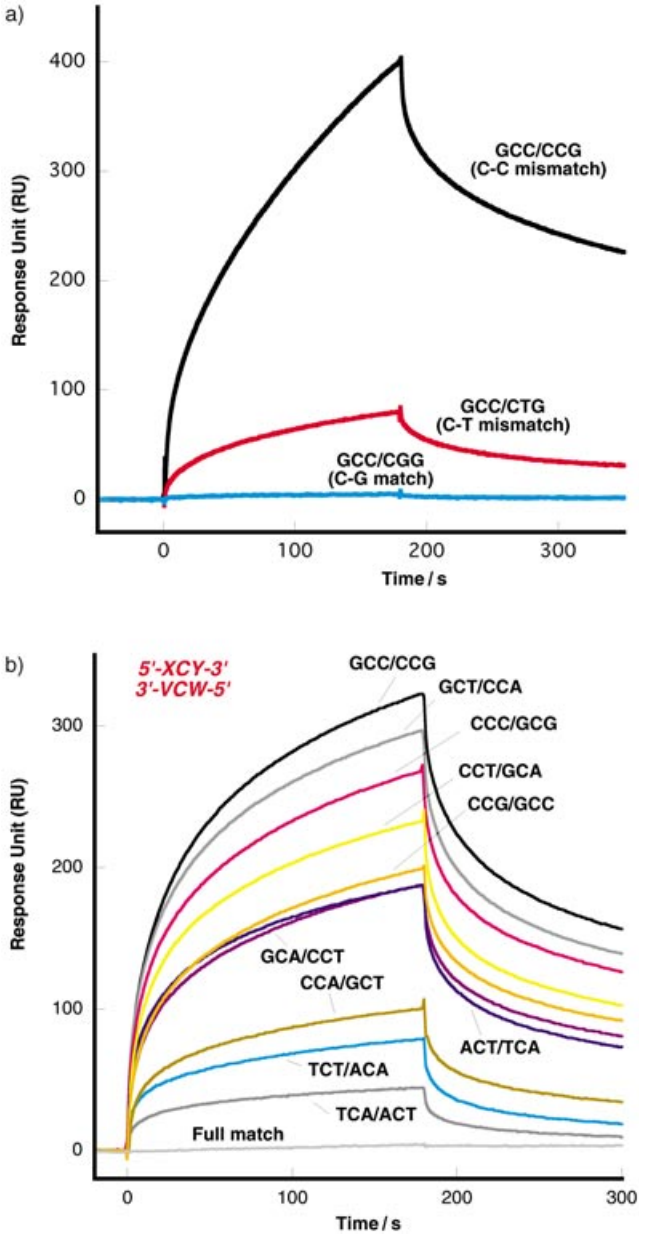


Figure 12. *SPR analyses of a) C–C and C–T mismatches containing 27-mer duplexes and b) C–C mismatch containing duplexes with* ***amND****-immobilized sensor.*

There are eight mismatched base pairs. It is desirable for MBL-based heteroduplex analyses to have eight different ligands that selectively bind to one mismatch. But, as shown in Figure 8, two heteroduplexes with different mismatches are simultaneously produced. Thus, at least four binding ligands are necessary for complete SNP typing. The molecular design of ligands targeting other mismatches is currently in progress in the laboratory.

5. Future Perspectives

How many SNPs should be typed for the diagnosis of a particular disease in the future? The number of SNPs is an important factor for the typing methods. A large-scale typing is suitable for the array format, whereas an homogeneous assay format would have an advantage for small-scale typing. We do not have an obvious answer yet. This is partly because a precise specification of the methods required for practical SNP typing is not obvious at this moment. But surely it depends on the type of disease; either common diseases or genetic diseases, polygenic or monogenic. Therefore, it is probable that not one but several typing methods will be used in different situations, depending on the practical demands, target diseases, and also size of the samples. Efforts in the chemistry area toward developing new and effective methods of nonenzymatic SNP typing are essential and should be continued. Among the many pro-

posed methods based on diverse concepts and principles, really useful and practical methods will be necessarily selected.

An important issue for all SNP typing, not only for enzyme-assisted but also for nonenzymatic methods, is the necessity of a large number of oligonucleotides for the assay. PCR amplification needs two primers. ASO hybridization needs another two hybridization probes. Therefore, for the typing of one SNP based on allele-specific hybridization, at least four oligonucleotides are needed. Assuming that tying of 100 SNPs was necessary for the particular disease, the number of oligonucleotides needed for the assay would be 400. As long as the assay used PCR amplification and oligonucleotide probes, we could not reduce this number and, consequently, the cost necessary for synthesis and a complicated handling of many oligonucleotides. The ultimate technology we should pursue is the direct detection of the nucleotide at the polymorphic site in the genome without amplification and usage of oligonucleotide probes. To achieve this goal, cooperation with scientists in other areas is essential. The author wishes that this short review would be an opportunity to begin a new collaborative work between chemists and scientists in other disciplines.

Keywords: alleles · DNA · mismatch-binding ligand · oligonucleotides · single nucleotide polymorphisms

[1] D. G. Wang, J. B. Fan, C. J. Siao, A. Berno, P. Young, R. Sapolsky, G. Ghandour, N. Perkins, E. Winchester, J. Spencer, L. Kruglyak, L. Stein, L. Hsie, T. Topaloglou, E. Hubbell, E. Robinson, M. Mittmann, M. S. Morris, N. P. Shen, D. Kilburn, J. Rioux, C. Nusbaum, S. Rozen, T. J. Hudson, R. Lipshutz, M. Chee, E. S. Lander, *Science* **1998**, *280*, 1077.
[2] A. J. Brookes, *Gene* **1999**, *234*, 177.
[3] A. J. Schafer, J. R. Hawkins, *Nat. Biotechnol.* **1998**, *16*, 33.
[4] J. J. McCarthy, R. Hilfiker, *Nat. Biotechnol.* **2000**, *18*, 505.
[5] P. Y. Kwok, *Annu. Rev. Genom. Hum. Genet.* **2001**, *2*, 235.
[6] A.-C. Syvänen, *Nat. Rev. Genet.* **2001**, *2*, 930.
[7] A.-C. Syvänen, *Hum. Mutat.* **1999**, *13*, 1.
[8] J. Tost, V. G. Gut, *Mass. Spectrom. Rev.* **2002**, *21*, 388.
[9] T. G. Drummond, M. G. Hill, J. K. Barton, *Nat. Biotechnol.* **2003**, *21*, 1192.
[10] R. W. Kwiatkowski, V. Lyamichev, M. de Arruda, B. Neri, *Mol. Diagn.* **1999**, *4*, 353.
[11] S. Kim, H. D. Ruparel, T. C. Gilliam, J. Ju, *Nat. Rev. Genet.* **2003**, *4*, 1001.
[12] J. Monforte, C. Becker, *Nat. Med.* **1997**, *3*, 360.
[13] S. Kim, J. R. Edwards, L. Deng, W. Chung, J. Ju, *Nucleic Acids Res.* **2002**, *30*, e85.
[14] K. J. Livak, S. J. A. Flood, J. Marmaro, W. Giusti, K. Deetz, *PCR Methods Appl.* **1995**, *4*, 357.
[15] R. K. Saiki, P. S. Walsh, C. H. Levenson, H. A. Erlich, *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 6230.
[16] I. V. Kutyavin, I. A. Afonina, A. Mills, V. V. Gorn, E. A. Lukhtanov, E. S. Belousov, M. J. Singer, D. K. Walburger, S. G. Lokhov, A. A. Gall, R. Dempcy, M. W. Reed, R. B. Meyer, J. Hedgpeth, *Nucleic Acids Res.* **2000**, *28*, 655.
[17] V. O. Tobe, S. L. Taylor, D. A. Nickerson, *Nucleic Acids Res.* **1996**, *24*, 3728.
[18] H. Baron, S. Fung, A. Aydin, S. Bahring, F. C. Luft, H. Schuster, *Nat. Biotechnol.* **1996**, *14*, 1279.
[19] P. Hardenbol, J. Banér, M. Jain, M. Nilsson, E. A. Namsaraev, G. A. Karlin-Neumann, H. Fakhrai-Rad, M. Ronaghi, T. D. Willis, U. Landegren, R. W. Davis, *Nat. Biotechnol.* **2003**, *21*, 673.
[20] W. M. Howell, M. Jobs, U. Gyllensten, A. J. Brookes, *Nat. Biotechnol.* **1999**, *17*, 87.
[21] P. L. Ross, K. Lee, P. Belgrader, *Anal. Chem.* **1997**, *69*, 4197.
[22] W. J. Kim, Y. Sato, T. Akaike, A. Maruyama, *Nat. Mater.* **2003**, *2*, 815.
[23] I. V. Yang, H. H. Thorp, *Anal. Chem.* **2001**, *73*, 5316.
[24] S. Takenaka, K. Yamashita, M. Takagi, Y. Uto, H. Kondo, *Anal. Chem.* **2000**, *72*, 1334.
[25] K. Yamashita, A. Takagi, M. Takagi, H. Kondo, Y. Ikeda, S. Takenaka, *Bioconjugate Chem.* **2002**, *13*, 1193.
[26] *Topics in Current Chemistry 236: Long-Range Charge Transfer in DNA 1*, (Ed.: G. B. Schuster), Springer, Berlin, **2004**.
[27] E. M. Boon, D. M. Ceres, T. G. Drummond, M. G. Hill, J. K. Barton, *Nat. Biotechnol.* **2000**, *18*, 1096.
[28] M. E. Núñez, K. T. Noyes, J. K. Barton, *Chem. Biol.* **2002**, *9*, 403.
[29] S. O. Kelly, E. M. Boon, J. K. Barton, N. M. Jackson, M. G. Hill, *Nucleic Acids Res.* **1999**, *27*, 4830.
[30] A. Yamane, *Nucleic Acids Res.* **2002**, *30*, e97.
[31] M. J. Rist, J. P. Marino, *Curr. Org. Chem.* **2002**, *6*, 775.
[32] M. E. Hawkins, W. Pfleiderer, A. Mazumder, Y. G. Pommier, F. M. Balis, *Nucleic Acids Res.* **1995**, *23*, 2872.
[33] M. E. Hawkins, F. M. Balis, *Nucleic Acids Res.* **2004**, *32*, e62.
[34] A. Okamoto, K. Tanaka, T. Fukuda, I. Saito, *J. Am. Chem. Soc.* **2003**, *125*, 9296.
[35] A. Okamoto, K. Kanatani, I. Saito, *J. Am. Chem. Soc.* **2004**, *126*, 4820.
[36] G. T. Hwang, Y. J. Seo, B. H. Kim, *J. Am. Chem. Soc.* **2004**, *126*, 6528.
[37] K. Nakatani, S. Sando, I. Saito, *Nat. Biotechnol.* **2001**, *19*, 51.
[38] K. Nakatani, S. Sando, I. Saito, *J. Am. Chem. Soc.* **2001**, *123*, 12650.
[39] K. Nakatani, S. Sando, I. Saito, *Bioorg. Med. Chem.* **2001**, *9*, 2381.
[40] S. Hagihara, H. Kumasawa, Y. Goto, G. Hayashi, I. Saito, K. Nakatani, *Nucleic Acids Res.* **2004**, *32*, 278.
[41] A. Kobori, S. Horie, H. Suda, I. Saito, K. Nakatani, *J. Am. Chem. Soc.* **2004**, *126*, 557.
[42] K. Nakatani, A. Kobori, H. Kumasawa, I. Saito, *Bioorg. Med. Chem. Lett.* **2004**, *14*, 1105.
[43] K. Nakatani, A. Kobri, H. Kumasawa, Y. Goto, I. Saito, *Bioorg. Med. Chem.* **2004**, *12*, 3117.
[44] B. A. Jackson, J. K. Barton, *J. Am. Chem. Soc.* **1997**, *119*, 12986.
[45] B. A. Jackson, V. Y. Alekseyev, J. K. Barton, *Biochemistry* **1999**, *38*, 4655.
[46] E. R. Lacy, K. K. Cox, W. D. Wilson, M. Lee, *Nucleic Acids Res.* **2002**, *30*, 1834.
[47] A. J. Nataraj, I. Olivos-Glander, N. Kusukawa, *Electrophoresis* **1999**, *20*, 1177.
[48] R. L. Rich, D. G. Myszka, *Curr. Opin. Biotech.* **2000**, *11*, 54.

Received: May 5, 2004
Early View Article
Published online on November 8, 2004

