MRC-Holland MLPA®

Designing synthetic MLPA probes and probemixes

Table of Contents

1. Introduction

Multiplex Ligation-dependent Probe Amplification, MLPA, is the gold standard for the detection of gene copy number changes. MRC Holland has designed hundreds of probemixes covering a wide variety of genes and genetic disorders with multiple applications. We encourage you to first search our extensive range of products to see if a product exists that fits your needs or e-mail us to see if one is in development before beginning synthetic probe design. However, if the gene or region you are interested in is not covered by any of our existing probemixes you can design your own synthetic MLPA probes. To facilitate this, we have created the following protocol that describes the design of synthetic MLPA probes. When designing probes for human genes, we recommend adding your synthetic probes to our SALSA MLPA P200 Reference-1 probemix or SALSA MLPA P300 Reference-2 probemix. These probemixes contain high-quality reference probes and MLPA control fragments with sufficient space for the addition of synthetic probes. More information can be found in *10.1 Adding probes to P200/P300 Reference probemix.*

A maximum of 21 probes (preferably 11) can be added to the P200-B1 probemix and a maximum of 15 probes (preferably 8) can be added to the P300-B1 probemix; an appendix of suitable probe lengths is available upon request and see section *5 Basics of MLPA probe design* and *6 Important design concerns*. If you choose not to use the P200 or P300 reference probemix, your MLPA reaction must contain at least 9 synthetic self-designed MLPA probes, of which 8 probes are reference probes. Using fewer probes makes the reaction unreliable.

Self-designed synthetic probes differ from MRC Holland probes in that synthetic probes have limited length possibilities. MRC Holland probes consist of one synthetic and one clone-derived oligonucleotide (oligo) allowing for the production of longer probes and the possibility to include up to 60 different probes in one MLPA reaction.

1.1. Advantages of self-designed synthetic probes vs. MRC Holland probes:

- Your specific region(s) of interest is targeted.
- Available quickly.

1.2. Disadvantages of self-designed synthetic probes vs. MRC Holland probes:

- Maximum probe length of 168 nt.
- Restrictions on maximum probe length due to limitations with oligo production quality.
- Fewer probes per MLPA reaction.
- Quality of probes depends on your design ability and oligo supplier.
- Lack of experience with probe design. This protocol aims to describe MLPA probe design as comprehensively and clearly as possible; however, it is impossible to convey all in-house probe design experience.
- Probe validation cannot be performed to the level that MRC Holland tests its probes. At MRC Holland probes are extensively tested for various characteristics under multiple conditions including peak height, reproducibility across samples, and sensitivity to differences in salt concentration and polymerase activity.

Feedback about this protocol? Please contact us via *support.mlpa.com* or *info@mrcholland.com*.

Warning

This protocol offers detailed guidelines for synthetic probe design. MRC-Holland cannot offer any additional support on synthetic probe design. If you are new to MLPA, please get familiar with the MLPA method by visiting *support.mlpa.com* and reading the MLPA General Protocol. MLPA is a patented technology and a license fee will be due in case you produce a self-designed MLPA probemix and offer it for sale to third parties, or in case you use self-designed MLPA probes in your own facility for commercial applications involving more than 5000 reactions / year.

2. Terminology

An MLPA probe consists of two oligonucleotides: a Left Probe Oligonucleotide (LPO) and a Right Probe Oligonucleotide (RPO) (**Figure 1**). Both the LPO and RPO may contain an optional stuffer sequence between the hybridising sequence and the primer binding site. If possible, stuffer sequences should be avoided in self-designed synthetic probes.

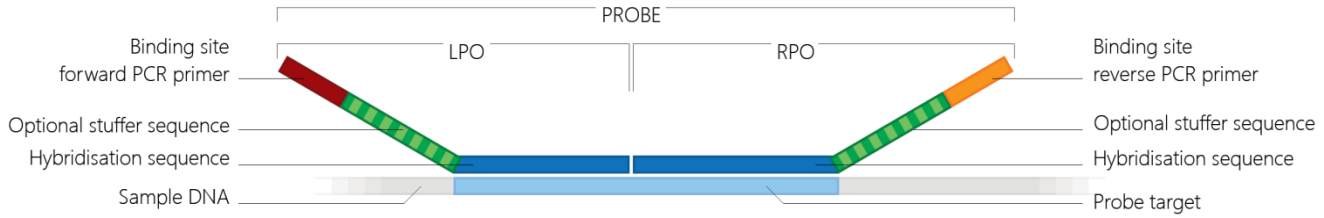


Figure 1 - Terminology of probe components. LPO: 5′ end: binding sequence of the forward PCR primer (GGGTTCCCTAAGGGTTGGA); 3′ end: Left Hybridising Sequence (LHS). **RPO**: 5′ end: Right Hybridising Sequence (RHS); 3′ end: binding sequence of the reverse PCR Primer (TCTAGATTGGATCTTGCTGGCAC). The primer binding sequences are used for the amplification of the probe during the PCR reaction; fluorescently labelled primers are available from MRC Holland.

Table 1 - Terms and abbreviations used in the designing synthetic MLPA probes and probemixes protocol

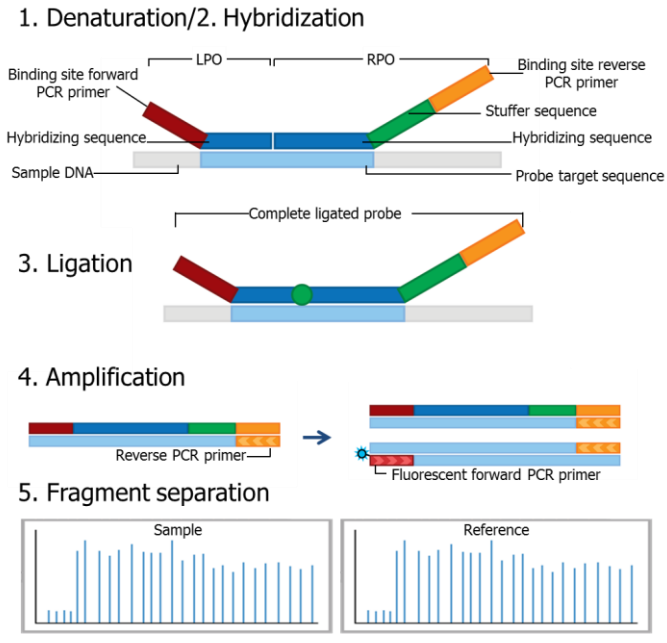
Term/abbreviation	Explanation
BLAST	Basic Local Alignment Search Tool (BLAST) from NCBI, used to compare a query DNA sequence with sequences deposited in the GenBank database. For BLAST searches in probe design, use the Human genomic + transcript database (sequence vs. "standard" human genome) and the Others (NR (Non-Redundant) etc.) database (sequence vs. collection of selected DNA/mRNA sequences). *https://blast.ncbi.nlm.nih.gov/Blast.cgi*
BLAT	BLAST-like alignment tool (BLAT). Developed by the University of California, Santa Cruz. BLAT is a tool used to find sequence similarities, flanking sequences, SNPs, CpG islands, copy number variations in healthy individuals, and many more features of a query sequence. *https://genome.ucsc.edu/cgi-bin/hgBlat?command=start*
Genome Data Viewer	NCBI website that offers a graphical overview of the human and other genomes: *https://www.ncbi.nlm.nih.gov/genome/gdv/*
HUGO name	Each gene has been assigned an official HUGO (Human Genome Organization) name by the HUGO Gene Nomenclature Committee. MRC Holland always refers to genes by their HUGO names. More information on *http://www.genenames.org*
Left Probe Oligo - LPO	Probe oligonucleotide that is situated on the left when probe is shown from 5′ to 3′. Consists of forward PCR primer sequence (5′ end), optional stuffer sequence, and hybridising sequence (3′ end).
Right Probe Oligo - RPO	Probe oligonucleotide that is situated on the right when probe is shown from 5′ to 3′. Consists of hybridising sequence (5′ end), optional stuffer sequence, and the reverse complement of the reverse PCR primer (3′ end). The 5′ end of the RPO **must** be phosphorylated or ligation of the LPO and RPO will fail.
Left Hybridising Sequence & Right Hybridising Sequence (LHS & RHS)	DNA hybridising parts of the LPO and RPO. The LHS and RHS bind to adjacent target DNA sequences.
Ligation site	3′ nucleotide of LPO and 5′ nucleotide of RPO, where ligation occurs between the two oligonucleotides.

Term/abbreviation	Explanation
Melting temperature - Tm	Temperature at which 50% of a sequence's copies are in a stable double helix and 50% are single stranded. The Tm indicates the strength of the probe-target binding and is influenced by the length and sequence (in particular the %GC) of the hybridising sequence, salt concentration, and the solvent used. RaW-Probe software is the tool used by MRC Holland to calculate the Tm; to download RaW-Probe visit *https://support.mlpa.com/downloads/files/raw-probe-program*. Please note that many algorithms exist to calculate Tm, and each gives a slightly different value.
NM_ sequence	Refers to the RefSeq database reference sequences of an mRNA. Each common transcript variant has a separate NM_ accession number. An overview of the various transcript variants of a gene can be obtained on the Entrez Gene website: *https://www.ncbi.nlm.nih.gov/gene*. For exon numbering, which is not part of the NM_sequence, look for the corresponding NG_sequence here: *http://www.ncbi.nlm.nih.gov/refseq/rsg/browse/*. NG_sequences also include intronic sequences.
nt	Nucleotide.
MLPA probe	The **combination** of the LPO and RPO.
Primer	DNA oligonucleotide that can anneal to a complementary DNA sequence and be used as the starting point for extension by a polymerase enzyme during PCR. A universal primer pair is used for all MLPA probes, including synthetic probes; fluorescently labelled PCR primers are supplied by MRC Holland as part of a SALSA MLPA EK reagent kit. **Note - in MLPA, the use of the word primer is not the same as a probe!**
SNP	Single Nucleotide Polymorphism.

3. MLPA method

3.1. MLPA method

MLPA is carried out in 5 steps, as depicted and described in **Figure 2**.



1. DNA sample is denatured by heating and a probemix is added.
2. Each probe consists of two parts: the LPO and RPO that hybridise to adjacent target DNA sites on the sample DNA.
3. LPOs and RPOs that hybridised to directly adjacent target sequences are ligated to form a complete probe.
4. A PCR reaction exponentially amplifies complete probes using a universal primer set.
5. Amplification products are separated by capillary electrophoresis. Analysis is performed using intra- and inter-normalisation of probe signals by the free MLPA analysis software (Coffalyser.Net) making it possible to compare a sample of interest with multiple reference samples. This analysis yields information on gene copy number.

Note that a cloned RPO (as shown here) always contains a stuffer sequence (green). In synthetic MLPA probes, the use of a stuffer sequence is not necessary or recommended.

Figure 2 - MLPA method. Visual representation and a description of each step of the MLPA reaction. For more details about the MLPA method visit the "Technology" section on the MLPA website: *www.mrcholland.com*, or visit the "Getting Started" section of *support.mlpa.com*.

3.2. Methylation-specific MLPA method

Methylation-specific MLPA (MS-MLPA) probes contain an HhaI restriction site in their hybridising sequence, but retain all the other elements of conventional MLPA probes (**Figure 3**). The HhaI enzyme cuts double stranded DNA at the GCGC motif except if the GCGC restriction site motif is methylated (CpG methylation). Each MS-MLPA reaction is divided in two parts after the probe hybridisation reaction: one for copy number analysis (ligation only) and one for the methylation status analysis (ligation + digestion). When the HhaI enzyme is added with the ligase-65 enzyme, probes that are hybridised to DNA with an unmethylated HhaI restriction site will be digested and will not be amplified in the subsequent PCR reaction step.

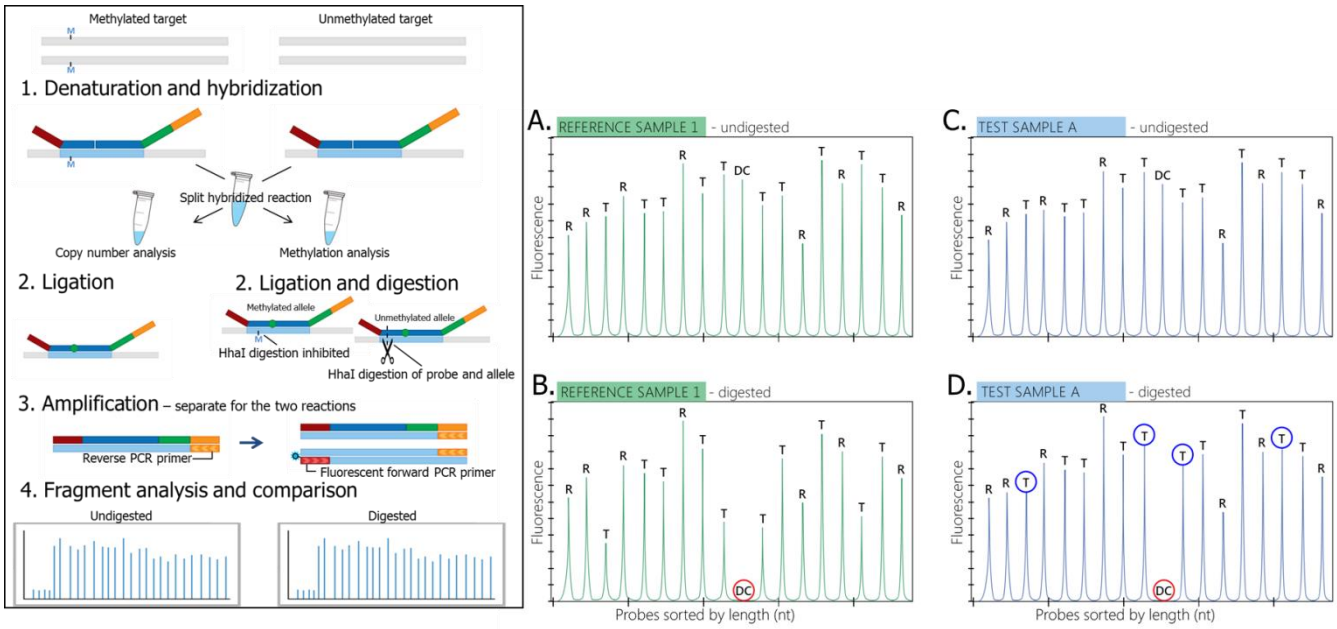


Figure 3 - Left: overview of the MS-MLPA method. After hybridisation, the reaction is split into an undigested reaction, for copy number analysis, and a digested reaction, for methylation analysis. In the digested reaction, the HhaI restriction enzyme will cut probes that are hybridised to unmethylated restriction sites; cut probes will not be amplified. **Right: results from an MS-MLPA experiment** A) Undigested reference sample, B) Digested reference sample, C) Undigested patient sample, D) Digested patient sample. Aberrant methylation is visible by comparison of pictures B and D where probes with aberrant methylation are indicated with a blue circle. The probe labelled DC and indicated with a red circle is a digestion control probe.

4. DNA sequence for MLPA probes

4.1. Getting the DNA sequence of interest

To design a probe, the target DNA sequence of interest is required. One way to find and format the sequence of a human gene is described below (*4.2 Formatting the DNA Sequence*). If the target DNA sequence of interest is already known, please read section *4.3 Elongating a known human DNA sequence in one or both directions*. For non-human organisms, other databases and tools may be better suited than those described below.

1. **Determine the name of the gene of interest.** Many genes have multiple aliases; to avoid confusion it is recommended to check *http://www.genenames.org/* to determine the HUGO-approved gene name.
2. **Find the accession number of the mRNA reference sequences.** Accession numbers can be found on NCBI Entrez gene: *https://www.ncbi.nlm.nih.gov/gene* or on the Genome Data Viewer (GDV) website: *https://www.ncbi.nlm.nih.gov/genome/gdv/*, search using the gene name. From the results, identify the associated RefSeq transcripts.
   - A list of Reference Sequences associated with the gene are shown, these are often splice variants. Find the NM_sequences and select the standard transcript that is used in the RefSeqGene project (it might be necessary to check all NM_sequences), or choose the longest transcript. Keep in mind that the

standard transcript used in the RefSeqGene project or the longest transcript does not necessarily include all exons. If this is the case, look for missing exons in alternative transcripts. Write down the NM_ accession number of choice for later use. The NM_sequence can also be searched for or checked using *https://www.lrg-sequence.org/index.html*, however, not all genes are present in this data base.

> **Exon numbering**: NG_sequences can be consulted to find the exact exon numbers. Keep in mind that intronic sequences are also included in the NG-sequences.

- If you use the GDV website, an overview is given of the genomic position of the gene relative to other genes. Click on the gene of interest again to expand and view available mRNA reference sequences, for an overview of steps see **Figure 4**.

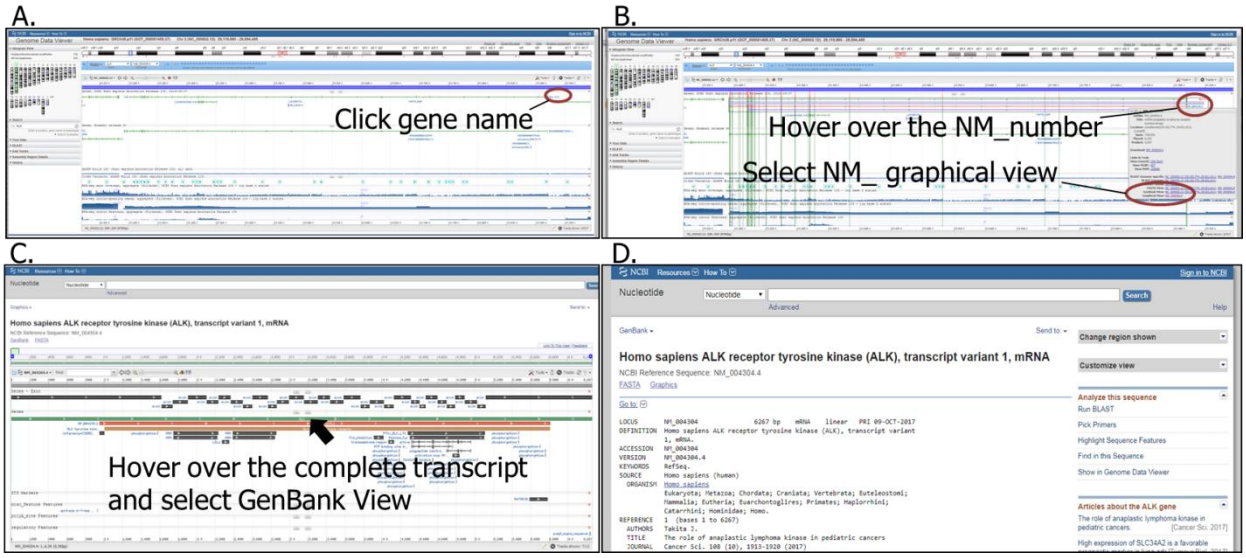


Figure 4 - **Steps required to obtain the DNA sequence.** A) GDV gene search results, B) Selecting NM_number C) NM_sequence graphical gene overview, D) NCBI Reference Sequence.

3. **Determine intron-exon boundaries.** To determine intron-exon boundaries, compare the mRNA sequence with the complete genomic sequence. One way to do this is via Genome BLAST on the NCBI website (*https://blast.ncbi.nlm.nih.gov/Blast.cgi*). Paste the sequence of interest or the GenBank (NM_) accession number and BLAST using the default settings. On the next screen, change the Alignment View to Query Anchored with dots for identities, and limit your results to "Homo Sapiens" in the Organism field.

4. **Retrieve Genomic Sequence.** Search your gene name of interest in the UCSC Genome Browser: *https://genome.ucsc.edu/*. Under the heading RefSeq Genes, select the desired RefSeq with corresponding accession number. Alternatively, submit the sequence of interest in BLAT (*http://genome.ucsc.edu/cgi-bin/hgBlat*). Click on the RefSeq/accession number of interest; an overview of the part of the genome where the gene of interest is located will be loaded. When the browser is open, select the following tracks to be displayed and then click refresh:

Mapping and Sequencing Tracks
- Base Position: full
- Chromosome Band: dense

Genes and Gene Prediction Tracks
- RefSeq Genes (NCBI RefSeq): pack

Regulation
- CpG Islands: show

Variation
- All SNPs (150, or most recent): dense

Repeats
- RepeatMasker: dense
- Simple Repeats: dense

Determine the direction of the gene (>>>> Forward; <<<< Reverse). Click View at the top of the screen and select DNA. In the subsequent window tick the option "All lower case". If the gene is in the reverse orientation as indicated on the previous screen, tick the *Reverse Complement* (get '-' strand sequence) box (see **Figure 5**). Click the button *extended Case/Color options*, copy the settings from **Figure 6** and click *submit*. (If some tracks are unavailable, go two pages back and change them from "hide" to "full". Click refresh to apply the settings.) Copy the sequence obtained to MS Word.

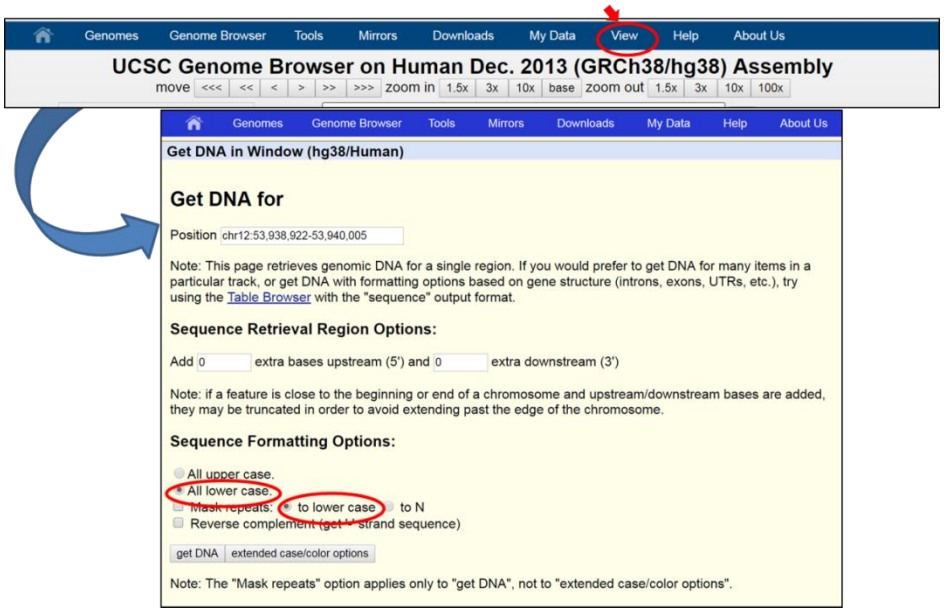


Figure 5 - Obtaining the DNA sequence from the Santa Cruz website.

Track Name	Toggle Case	Under-line	Bold	Italic	Red	Green	Blue
CpG Islands	☐	☐	☐	☐	0	0	255
RepeatMasker	☐	☐	☐	☐	0	0	0
NCBI RefSeq	☑	☑	☐	☐	0	0	0
All SNPs(150)	☐	☐	☐	☑	0	200	0
Chromosome Band (Ideogram)	☐	☐	☐	☐	0	0	0
Simple Repeats	☐	☐	☐	☐	0	0	0

Figure 6 - "Extended Case/Color options" in UCSC Genome Browser DNA retrieval window.

4.2. Formatting the DNA Sequence

The steps below describe formatting DNA Sequences in Word, you can also use any other sequence annotation program of choice.

1. Remove all line breaks/enters using the *Find and replace* tool in Word.
2. Click Ctrl+H in Word to open it. Next, type *^p* in the *Find what:* field. Leave the *Replace with:* field empty. Click *Replace All*.
3. To make SNPs more visible it is useful to highlight them. In the *Find and replace* tool click the button *More >>*; see **Figure 7**. Place the cursor in the empty *Find what:* field, click *Format, Font choice*, select font style Italic and click *OK*. This searches for everything in the document that is in italics, these are the SNPs. Now place the cursor in the empty *Replace with:* field. Click *Format*, select Highlight. The colour of the highlight can be changed using this button  in the toolbar of the document. Click *Replace All*. Please note that it may be impossible to avoid all SNPs during probe design. Please see section *6.4.1.1 Designing probes to minimise the effect of a mismatch* for more information.

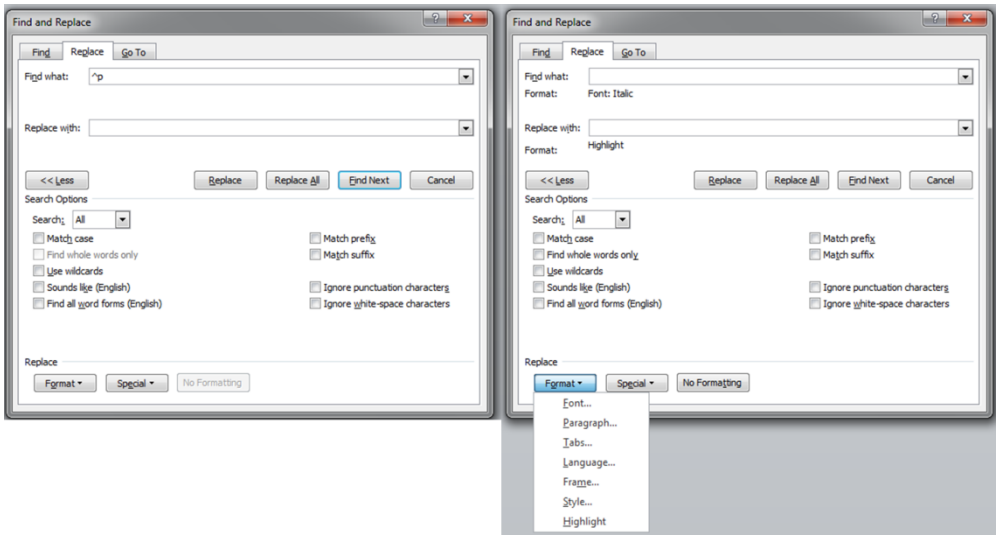


Figure 7 - "Find and Replace" function in Word with expanded options by selecting More>>.

4. When designing methylation probes for MS-MLPA, it is useful to indicate the restriction sites in your sequence. When using HhaI, this is GCGC. In the *Find and replace* tool set the cursor in the *Find what:* field, type the restriction site sequence and make sure all formatting restrictions are off by clicking *No Formatting*. Type the restriction site in the *Replace with:* field, click *Format*, select a different Highlight colour than you previously used, and click *Replace All*. The result should be similar to that shown in **Figure 8**.

AGGGGTGGAGAGAAAAGAGGGGAGGGATGGGGGGAGGGGAAACAGGAGCGAGGTGTCTCCCTAGCTCGCTGCCTCTGGCAAGTGGAGT
TTTTAAAAAGCTCCAGCAGATCATGTCATGACGACTTCGCTGCTCCTGCATCCACGCTGGCCGGAGAGCCTTATGTACGTCTATGAGG
ACAGCGCGGCGGAGAGCGGCATCGGCGGCGGCGGCGGAGGAGGAGGCGGCGGCACGGGCGGAGCGGGGGGTGGCTGCAGCGGAGCGAG
CCCCGGCAAAGCCCCGAGCATGGATGGTCTGGGCAGCAGCTGCCCGGCCAGCCACTGCCGCGACCTGCTTCCGCACCCCGTGCTGGGC
CGCCCGCCGGCTCCCCTGGGCGCCCCTCAGGGCGCCGTCTATACGGACATCCCGGCCCCGGAGGCGGCGCGCCAGTGTGCCCCGCCGC
CCGCACCCCCCACCTCGTCCAGCGCCACCCTGGGCTACGGCTACCCCTTCGGGGGCAGCTACTACGGCTGCCGCCTGTCGCACAACGT
GAACCTGCAGCAGAAGCCTTGCGCCTACCACCCGGGCGATAAATACCCGGAGCCGTCGGGCGCCCCTGCCCGGTGACGACCTGTCCTCT
AGGGCCAAGGAGTTCGCCTTCTACCCCAGCTTCGCCAGCTCCTACCAGGCGATGCCCGGCTACCTGGACGTGTCGGTGGTGCCCGGGA
TCAGCGGGCACCCGGAGCCGCGTCACGACGCCCTCATCCCCGTCGAAGGCTACCAGCACTGGGCTCTCTCCAATGGCTGGGACAGTCA
GGTGTACTGCTCCAAGGAGCAGTCGCAGTCCGCCCACCTCTGGAAGTCTCCCTTCCCAGgtaaggaagggacccgagcgccgccgccg
ccggggaccctcccgcctgcctgccccggggctccgcgcccaaccacccccgccgtctggccccgcgcgcccgctcggctgggg
ctgcctatggagccggccgggcgagctgcactgaggaatgcgccggggaagaaatctgctccgacacgttctctgtagctgcccggcc
gagaatgaagcaatcacaggcgccccgaaagccgggccgccggctctgctctgtccggtagctcgcctccgcctcccctgcaggctcc
agcctcccgccggctcttggcccctaaacctgcttccggcaagggatgggggcggggtgggcctatagtgccttggaatccaggaca
aaacccccaacccaccgaataactggggagggcggagaataagaacccccactttctttgacagaattcgcaggatcgttcaggcact
agacagtatttttaaatagggactaatttgctgggtctacagaaatgtgagatttatttttttcctttcctgacttatttaaaa

Figure 8 - Formatted DNA sequence. Highlighted restriction sites and SNPs. Underlined and Uppercase: exon. Lowercase: intron. Blue: CpG island. Highlighted in light blue: HhaI restriction site (for methylation probes). Highlighted in brown and light blue italic: SNPs (avoid).

4.3. Elongating a known human DNA sequence in one or both directions

When only a part of a human DNA sequence is known (for instance the sequence detected by one of our probes), use the UCSC genome browser to retrieve the flanking genomic sequences.

- Paste the available DNA sequence in the UCSC genome browser BLAT Search Genome tool: *http://genome.ucsc.edu/cgi-bin/hgBlat*.
- Click *submit*.
- Select the sequence with the best homology, usually the top listed hit. Make note of the strand orientation, "+" or "-", then click browser.
- In the next screen, click View at the top of the screen and select DNA.

- Fill in the desired number of extra nucleotides upstream and downstream. Use multiples of 50 nt as each line of the sequence obtained is 50 nt. If the strand was "-", click *Reverse complement*. When additional information is required such as the position of SNPs, click *All lower case* and *extended case/color options* and proceed as depicted in **Figure 6**.

5. Basics of MLPA probe design

5.1. General probe design rules

Please also follow the guidelines in section *6 Important design concerns*. When designing probes to detect methylation, also follow the guidelines in section *7 Designing methylation-specific MLPA probes*.

- LPO Hybridising Sequence (LHS) and RPO Hybridising Sequence (RHS) should be located directly adjacent to each other.
- Hybridising sequences should ideally be located in the coding regions of a gene, the exonic sequence between the start and stop codon where more information on SNPs is available.
- All probes should have unique hybridising sequences (no overlap between probes).
- Each probe must have a unique total probe length (LPO+RPO).
- Total probe length mentioned in this protocol consists of: LHS + RHS + 42 nt (the length of both PCR primer binding sequences), and the length of a stuffer sequence if used (only use if necessary, see section *6.2 The effect of the Tm value*).
- Divide the total probe length evenly over the LPO and RPO to avoid an unnecessarily long oligo (note: longer oligos tend to be of lower quality).
- Minimum length of LHS **or** RHS: **21 nucleotides, preferably 26-30 nucleotides.**
- Minimum total length difference required between probes: **4 nt.**
- Total minimum and maximum synthetic probe length.
    - The ideal length for synthetic probes is between: **100 and 140 nt**.
    - The exact minimum and maximum probe length you design depends on the probemix you will add your synthetic probes to. Check the product description of the probemix you are going to add your probes to and make sure there is at least 4 nt between your probe(s) and any existing probe or control fragment in the probemix.
    - The maximum length for synthetic probes is **168 nt**. Use stuffer sequences only if absolutely necessary, see *6.2 The effect of the Tm value*. Long synthetic probes can have decreased signals due to the decreased quality of long synthetic oligos, and will often have shoulder peaks that can complicate data analysis. For this reason, we strongly recommend **avoiding the use of synthetic probes with a total length > 140 nt**.
- Tm of each hybridising sequence (LHS/RHS separately) should be **≥ 68°C** (preferably **≥ 71°C** *see* **Table 2**).
- ΔG = secondary structure (LHS / RHS) preferably should be **≥ 0** (*if not feasible drop this criteria*).
- LPO primer binding sequence[1]: **GGGTTCCCTAAGGGTTGGA.**
- RPO primer binding sequence[2]: **TCTAGATTGGATCTTGCTGGCAC.**
- GC content should be **~ 50%** (*If possible*).
- LHS : use a maximum of 2 G/C nts in the last 5 nt at the 3' end, directly adjacent to the ligation site, if possible.
- LHS/RHS: use a maximum of 3 consecutive G/C nts directly adjacent to the primer binding sequence, if possible.
- Avoid stretches of single nucleotides within the LHS and RHS as best as possible.
- When designing probes for consecutive exons they should not be successive in total probe length. (e.g. exon 1 - 100 nt, exon 2 - 104 nt).

[1] The primer sequence contained in the LPO is identical to the sequence of the forward PCR primer.
[2] The primer sequence contained in the RPO is complementary to the reverse PCR primer.

Table 2 - LHS/RHS length and Tm

Length (nt)	Tm range
21-25 (only use 21 nt - 24 nt in exceptional cases)	≥ 72.5°C, ideally > 74.0°C
26-30	≥ 71.0°C, ideally > 72.5°C
31-35	≥ 71.0°C
36-40	≥ 70.0°C
40-55	≥ 68.0°C

5.2. Probe design steps

- Find the region of interest, preferably in an exon. Try to find a region with the following features:
  - GC content = ~ 50%. Especially important for the 5' and 3' ends of the LHS.
  - No or limited homology with other human sequences. In the case of two closely related sequences, make certain that the last nucleotide (3' end) of the LPO has a mismatch with the related sequence *(6.4 The effect of mismatches, including SNPs and mutations, on probe signal*).
  - No SNPs. See section *6.4 The effect of mismatches, including SNPs and mutations, on probe signal.*
  - No overlap with another probe, as it will compete for the same target sequence.
- Adjust the length of the LHS and the RHS so that:
  - Tm is ≥71°C (see **Table 2**). Use RaW-Probe program[3].
  - The LHS 5' does not start with an Adenine (or the stuffer if one is used).
  - The length of the LHS and the RHS are at least 23 nucleotides each.
- Use the *UNAfold-Mfold-DNA Folding Form* to test the ΔG of the LPO and RPO:
  - Copy and paste the primer binding sequence, (stuffer) and LHS into the online form in that order.
  - Set the *[Na+]* = 0.35 M.
  - Set the *Folding temperature* to 60°C.
  - Click the *Fold DNA* button, ΔG should be ≥ 0. Always take the first ΔG value (Structure 1).
  - You can click the 'jpg'-link to see the folding structure of the oligo. This gives an indication as to the chances of the oligo binding to itself, especially in the case of a negative ΔG.
  - Repeat the steps above for the RPO (note; the LPO and RPO have a different primer sequence).
- Checking probe specificity:
  - Check the LHS and RHS for alignment to all known human sequences using NCBI's human genome + transcript database BLAST tool, see section *14 Useful websites, tools and software*. Select *somewhat similar sequences* in the *Program Selection* section.
  - If homologies exist, outside your intended target region, ensure that the LHS and RHS do not **each** have an overlap of >12 nt and a Tm >50°C (use RaW-Probe to calculate) with the homologous sequence. If **both** the LHS and RHS have an overlapping sequence region consisting of >12 overlapping nts **with** a Tm >50°C with a non-targeted sequence, both oligonucleotides may bind sufficiently to be ligated and generate a signal. If probe overlap with a homologous sequence is >12 nt and has a Tm >50°C with either the LHS OR the RHS but **not both** then this probe is still OK. See **Figure 9** for examples of non-targeted homologous sequences that overlap with a designed probe.
- Check probe specificity in the NR database BLAST: if the origin of the sequence is unclear, copy the name given under *LOCUS* and look it up in GDV and/or do a BLAT search via the UCSC genome browser.
- If a probe sequence is polymorphic or can be found in multiple genes design a new probe: see *6.4 The effect of mismatches, including SNPs and mutations, on probe signal*.

[3] The RaW-Probe program that can be downloaded from our website, use the following settings: Go-Oli-Go method; 0.1 M salt; 1 µM oligo concentration. Different programs will give different Tm values!

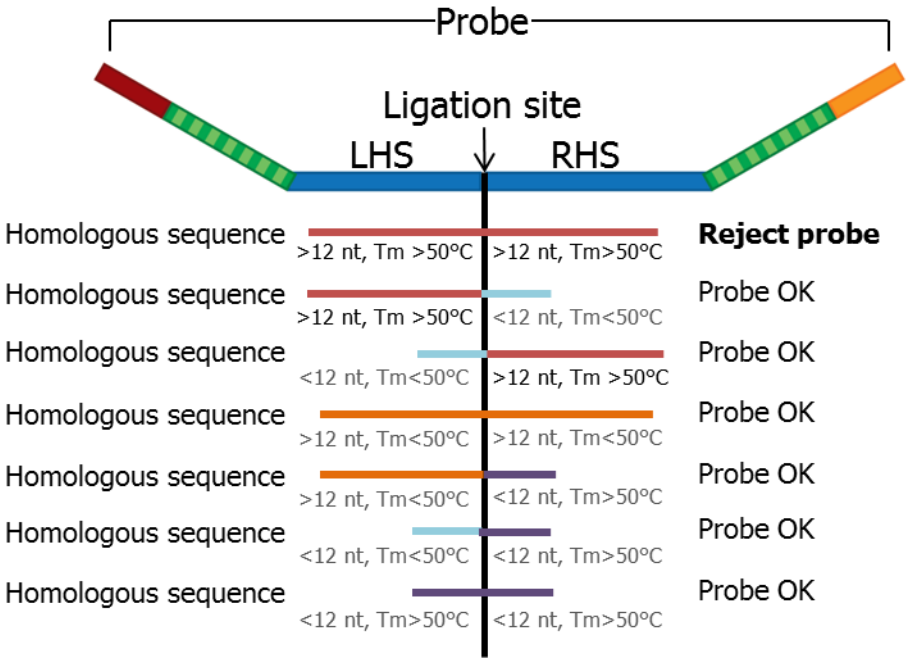


Figure 9 - Depiction of potential non-targeted homologous sequence overlaps with a self-designed synthetic probe. This image outlines a number of potential situations that may occur when designing synthetic probes and a portion of the probe overlaps with a non-targeted homologous sequence. This image can act as a guide as to when to accept a probe that contains sequence homologies to regions outside of your target region.

6. Important design concerns

6.1. Minimum number of MLPA probes per reaction

The quantity of MLPA probes affects the MLPA PCR reaction. In an MLPA reaction, all probes are amplified by the same primer set. When a sufficient number of probes are used the PCR reaction will stop due to primer depletion after ~30 cycles, when all PCR primer is incorporated into the amplification products. Once PCR primers are depleted from a reaction extra PCR cycles will not influence the relative peak heights/areas. In contrast, when a small number of probes (<5) is used, PCR primers will not be completely consumed, meaning the PCR reaction does not come to a full stop. As a result, relative peak heights can change in the last PCR cycles, thereby making the results more variable. Results are best when 9 probes are used. Please note we strongly advise that you add synthetic self-designed probes to the P200 or P300 probemix. Not only will you avoid issues with the number of probes in the mix, but you will also benefit from the stable reference probes present in each mix.

6.2. The effect of the Tm value

There is no upper limit for the Tm of a probe, making it possible to use long hybridising sequences. The advantage of a long hybridising sequence and a high Tm is that the probe signal will be less sensitive to polymorphisms (SNPs) within the hybridising sequences. The use of a non-hybridising stuffer sequence (derived from e.g. T7 or λ-phage sequences) may occasionally be advantageous to obtain sufficient probe length, for instance when the target sequence is very GC-rich. However, we recommend that you only use a stuffer sequence when necessary, as it is generally more advantageous to extend the hybridising sequence.

Once a probe binds during the MLPA hybridization step, the chance of a probe detaching from its target during the remaining portion of the hybridisation reaction is very low if the Tm is sufficient. If the binding of one of the probe oligonucleotides to the DNA template is not stable, an equilibrium will be reached between bound probe and unbound probe. This equilibrium is extremely sensitive to small changes in probe amount, incubation temperature, and salt concentrations. In these instances, small differences in evaporation between samples can also have a strong effect on the MLPA results. This is why we recommend having a Tm ≥71°C (see **Table 2**).

6.3. The effect of the first nucleotide

The first nucleotide following the LPO primer binding sequence (GGGTTCCCTAAGGGTTGGA**N**) affects the probe signal height. Usually this nucleotide will be the first nucleotide of the LPO hybridising sequence. When a stuffer sequence is used, this nucleotide will be the first nucleotide of the stuffer sequence. Nucleotide strength from strongest → weakest signal:

C > G > T > A.

The use of an Adenosine (A) as a first nucleotide should be avoided if possible. We recommend the following first nucleotides based on probe length:

- T for short probes (<120 nt)
- G for the intermediate probes (120-140 nt)
- C for long probes (>140 nt)

Using CC as the first two nucleotides will give an even higher probe signal (CC>CG>CT>CA). Note that the first nucleotide of choice does not necessarily need to hybridise to the target DNA and can be added as a non-hybridising nucleotide.

These recommendations may help reduce sloping, the effect that longer probes give lower signals, but are not essential for probe function. The above guidelines can also be used to redesign LPOs during MLPA experiment optimisation of the peak pattern to increase or decrease the peak height of selected probes. Please ensure that a minimum size gap of 4 nt between probes remains after the addition of any non-hybridising nucleotides.

6.4. The effect of mismatches, including SNPs and mutations, on probe signal

When designing a probe that includes an unavoidable SNP, or conversely, a probe that should specifically detect a single nucleotide difference such as a known point mutation, keep the following in mind:

The ligase-65 enzyme is most sensitive to mismatches at the 3' end of the LPO (LPO side of the ligation site).

Probes with a mismatch at the RPO side of the ligation site (5' end of the RPO) can be ligated to some extent. When mismatches are further away from the ligation site, the effect is less predictable with the mismatch potentially having no effect or causing a significant reduction in probe signal. This generally depends on the Tm of the remaining continuous sequence. In the next two sections more information is given on: how to design a probe that detects a mismatch, and how to design a probe to minimize the effects of a mismatch.

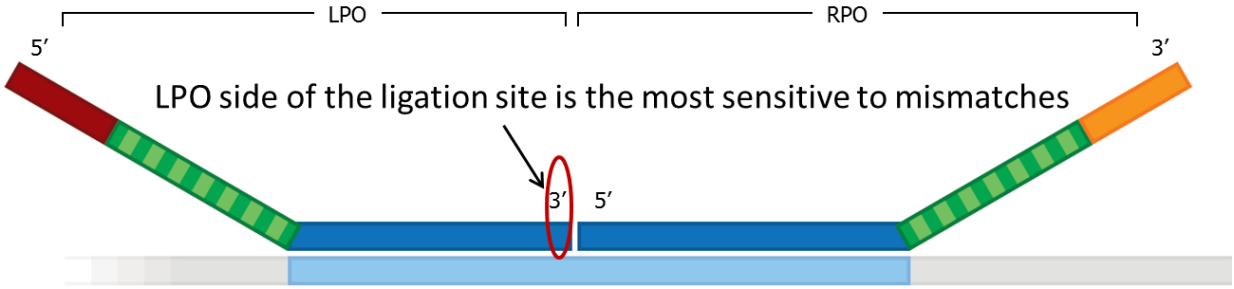


Figure 10 - The 3' end the LPO (the LPO side of the ligation site) is most sensitive to mismatches.

6.4.1. Designing probes to detect a mismatch

To design a probe that can distinguish two closely-related sequences (e.g. a gene from a pseudogene) or one that specifically detects a known point mutation/SNP, make sure the mismatch between the desired sequence to be detected and the related sequence or the point mutation/SNP is located on the last nucleotide on the 3' end of the LPO (LPO side of the ligation site); see **Figure 10**.

The probe you design will detect the complementary sequence (e.g. G in a probe will detect a C in the target sequence). If there is a mismatch with the probe at the 3' end of the LPO that results in a G/T or T/G pairing, the probe will still generate a signal that is approximately 25% of the signal without the mismatch. This is due to the fact that Guanine and Thymine are able to form two hydrogen bonds, allowing some ligation activity. A very small probe signal (<5%) may also be obtained when the probe has a C/A pairing with the sample DNA at the 3' end of the LPO.

Below, **Table 3** shows all possible pairings of the last base at the 3' end of the LPO and the expected probe signal generated from the pairing. This table assumes perfect base pairing with all other nucleotides of the probe (LHS and RHS).

Table 3 - Expected probe signals for pairing of 3' LPO

Probe sequence 3' end of LPO	Sequence paired with in the complementary strand of the target			
T	A 100% signal	G 25% signal	C no signal	T no signal
G	C 100% signal	T 25% signal	A no signal	G no signal
C	G 100% signal	A <5% signal	T no signal	C no signal
A	T 100% signal	G no signal	C no signal	A no signal

6.4.1.1. Designing probes to minimise the effect of a mismatch

Sequence variation within two nucleotides of the ligation site can result in a reduced probe signal due to a reduction in ligation efficiency of the two probe oligonucleotides. Mismatches in the middle of a hybridising sequence may also have an effect on the probe signal by negatively affecting probe binding stability. Therefore, any SNPs in the target sequence should be avoided, if possible.

If a known SNP cannot be avoided in the probe sequence, the effect can be minimised by ensuring the SNP is *at least 8 nt* from the ligation site. In addition, both the LHS and RHS sequence should have a remaining continuous sequence (without SNPs) with a Tm ≥70°C. This is to ensure that the probe oligonucleotide can bind stably to its target, despite the SNP presence in the DNA target sequence.

6.5. Selecting reference probes (only when it is NOT possible to use the P200 or P300 reference probemix)

MLPA is a relative technique, meaning that all obtained probe ratios are dependent on the other probes and samples within an experiment. To detect relative changes in peak signal generated by your target-specific probes, reference probes are needed. You should have multiple reference probes each targeting different chromosomes and they should detect genes that are unrelated to the condition of interest.

When you are working with human genomic DNA, adding your synthetic probes to the SALSA MLPA reference probemix P200 or P300 will provide you with sufficient reference probes and no additional reference probes need to be designed. If P200 or P300 cannot be used for human genomic DNA, then a minimum of 8 (for tumour characterization a minimum of 12) reference probes should be designed, all targeting different chromosomes. The design of reference probes follows the same rules as target-specific probes. You should select genomic regions that are not known to be affected in the condition of interest and where copy number changes are rare. Self-designed reference probes should also be equally distributed in length among all designed probes (e.g. reference probes should not all be the shortest or longest probes in your probemix). It is advised that your shortest and longest probe are reference probes.

For all non-human organisms, reference probes must be self-designed and a recommendation on a suitable number of reference probes cannot be given.

6.6. Optional signal-reducing competitor oligo

It is important that the peak height of all designed synthetic probes and the reference probes of a probemix are of a similar height to avoid problems associated with probe signal sloping. Once a probe has been designed, it should be tested within the desired probemix on commercially available DNA. Based on the results of this test a decision can be made if a competitor oligo (COMP) should be included in the probemix. The COMP is an oligo that is identical to the LPO and a small part (4 nucleotides: TGGA) of the forward primer binding sequence. The COMP will compete with the LPO for the limited number of binding sites on the genomic DNA within a DNA sample. The COMP can be ligated to the RPO, but the resulting COMP-RPO ligation product will not be exponentially amplified during PCR as the COMP-RPO probe only contains a full reverse primer binding sequence. By including a COMP in the probemix, the signal of a probe can be reduced. In general, the use of a 1:1 ratio of LPO and its corresponding COMP will reduce the probe signal by approximately 50%. A COMP oligo can be added together with the LPO and RPO oligos in the 200µl basic synthetic probemix, see *11.4 Making a competitor oligo mix.*

7. Designing methylation-specific MLPA probes

Please read the above sections: *5 Basics of MLPA probe design*, and *6 Important design concerns* carefully, these rules also apply to designing methylation-specific probes.

Cytosine residues followed by guanines (5' CpG 3') are targets for methylation enzymes in humans. CpG islands in the promoter region of genes are usually of interest for methylation testing. At MRC Holland the HhaI restriction enzyme, which digests probe-DNA hybrids containing an HhaI restriction site where target DNA is unmethylated, is used for detecting the methylation status of a specific DNA sequence. Any sequence harbouring an HhaI restriction site can, in theory, be used for MLPA methylation detection probes. It is also possible to design probes that use a different methylation sensitive endonuclease, however, not all enzymes will be compatible with the buffers and temperatures used in MLPA.

Restriction endonucleases are less efficient in digesting sites that are located near the end of a double-stranded region, in the case of MLPA the probe-target hybrid. Therefore, an HhaI site should not be located at the 5' end of the LHS or the 3' end of the RHS; make sure that there are at least 5 nt on either side of the GCGC HhaI restriction sequence that can hybridise to the target sequence to ensure efficient digestion.

- A GCGC motif, the HhaI restriction site, should be located within the hybridising sequence, with a minimum of 5 nts to the 5' end of the LHS or the 3' end of the RHS.
- Due to the high GC% of CpG islands, it can be difficult for all requirements specified under *5.1 General probe design rules* to be met. The following requirements can be adjusted:
    - ΔG can be slightly negative.
    - GC% should preferably be ~50% around the 5' and 3' of the LHS, but may be higher in other parts of the LHS.
- The majority of CpG sequences outside of promoter CpG islands are methylated in human DNA. CpG sequences within promoter CpG islands tend to be hypomethylated as H3K4me3 found at promoters/ transcriptional start sites repels DNA methylation. Sequences located near the boundaries of a CpG island may in some cases be methylated in a only subset of the cells, resulting in higher signals than expected for MS-MLPA probes directed to these sequences.
- For probes within CpG islands, a higher (2-3 fold) concentration of the probe oligonucleotides may be required as CpG islands have a tendency to form secondary structures. Secondary structures of a probe target sequence may reduce the oligonucleotides' hybridisation speed.

Please note that probes containing multiple HhaI sites will be digested if one of the GCGC motifs is unmethylated. Other criteria for MS-MLPA probes are similar to those of non-MS MLPA DNA probes (see *5 Basics of MLPA probe design*).

8. MLPA probe design example

Synthetic (MS-)MLPA DNA probe

Below is an example of a synthetic probe designed to the EVA1C gene. This probe also contains an HhaI restriction site that can be used for methylation determination of the region in an MS-MLPA experiment.

- MLPA probe for the human EVA1C gene.
- Chr. 21q22.11
- GenBank Sequence: NM_058187.4
- Total length of amplification product: (LPO) 54 + (RPO) 54 = 108 nt.
- **LPO**:
  - Tm of LHS= 77.2°C.
  - **forward primer sequence (bold)** + LHS:
  - **GGGTTCCCTAAGGGTTGGA**GGTTCAGAAAGATGCTGTGGCCCACTTTAAAACAA
  - Length: 19 (PCR primer; bold) + 35 (hybridising; underlined) = 54 nt.
- **RPO**:
  - Tm of RHS= 79.5°C
  - RHS (5' phosphorylated!!) + **reverse primer sequence (bold)**:
  - AGCCCAATTATTA***GCGC***TCGGCGGCTGTTTG**TCTAGATTGGATCTTGCTGGCAC**
  - Length: 31 (hybridising; underlined) + 23 (PCR primer; bold) = 54 nt.
  - HhaI site (GCGC) is highlighted in grey and italics.

The combined LHS+RHS is:

GGTTCAGAAAGATGCTGTGGCCCACTTTAAAACAAAGCCCAATTATTA***GCGC***TCGGCGGCTGTTTG

The probe will bind the complement of this sequence!

9. Ordering synthetic probes

We strongly recommend ordering probe oligonucleotides from IDT: *www.idtdna.com*. Low quality oligonucleotides (incomplete phosphorylation, length discrepancies) can result in lower peak signals and shoulder peaks. Specifically, an extra peak that is one nt shorter than the real probe peak can appear. Quality requirements for MLPA probe oligonucleotides are higher than for conventional PCR primers. For some probes that give a low signal, ordering a HPLC-purified LPO leads to an increase in probe signal. Purification of the RPO is not needed.

Before ordering ensure:

- All RPOs select 5'modification: 5' Phosphorylation.
- LPO: starts with GGGTTCCCTAAGGGTTGGA (forward primer binding sequence).
- RPO: ends with TCTAGATTGGATCTTGCTGGCAC (reverse primer binding sequence).
- Oligos shorter than 60 nt: 25 nmol.
- Oligos longer than 60 nt: "Ultramers" ultramer 4 nmol.

*Note: oligos should **not** be fluorescently labelled; only the PCR primers (ordered from MRC Holland) need to be labelled.

An example of a synthetic probe as it should be ordered can be found in paragraph *8 MLPA probe design example*. In our experience, synthetic oligonucleotides, including phosphorylated oligos, are very stable when dissolved in TE and stored at -20°C.

10. Ordering MLPA reagents

To perform an MLPA reaction, you need MLPA reagents. Reagents are sold by MRC Holland in SALSA MLPA EK kits; these contain all necessary enzymes, buffers, and labelled PCR primers. MLPA Reagent kits are described in detail in section *10.3 EK MLPA reagent kits (for standard MLPA and MS-MLPA).*

It is recommended that self-designed probes are added to a SALSA MLPA P200 or P300 Reference probemix, see *10.1 Adding probes to P200/P300 Reference probemix*.

More information about these options can be found below. All prices and ordering information can be found on *www.mrcholland.com* under Ordering > Price List & Payment.

10.1. Adding probes to P200/P300 Reference probemix

The SALSA MLPA P200 and P300 MLPA Reference probemixes contain carefully selected reference probes for human DNA and are specifically designed to be used in combination with self-designed synthetic human DNA probes. In addition, they contain various control fragments that help detect problems that could affect the MLPA reaction (insufficient DNA quantity, denaturation problems). Together, the reference probes and control fragments facilitate data analysis, maximise the number of synthetic probes you can include, and give you extra assurance that the MLPA reaction was performed correctly.

More details about the P200 and P300 Reference probemixes can be found on our website (*http://www.mrcholland.com* , search for P200 or P300). The P200 probemix leaves the size range from 88-170 nt open for the inclusion of synthetic probes, while P300 has reference probes distributed over the entire size range of the probemix including the smaller size range left open in P200. The latter allows for a better correction of signal sloping in the MLPA amplification products.[4] The use of P300 is recommended when a smaller number of synthetic probes are needed. The use of P200 is recommended when a large number of synthetic probes are needed. Advantages (+) & Disadvantages (-) of using P200 or P300 reference probemix:

+ No need to design reference probes. The SALSA MLPA reference probes in the P200 and P300 have been carefully selected from our probe database on the basis of their stable and reliable performance.
+ SALSA MLPA reference probes are located over the entire MLPA probe size range (P300) or predominantly in the longer probe range (P200), thus maximising the number of (shorter) self-made target probes that can be designed in the limited size range suitable for synthetic probes, without having to use up this design space for reference probes.
+ The presence of quantity control fragments (Q-fragments) warns for insufficient DNA or ligation problems.
+ Denaturation fragments (D-fragments) warn for poor DNA denaturation, for instance due to contaminates in the sample DNA.[5]
+ Digestion control probes for HhaI digestion indicate if digestion was complete when performing MS-MLPA.
+ P200 and P300 allow for easy MLPA troubleshooting: these probemixes enable you to determine whether possible problems are due to the synthetic probes (i.e. design, quality), sample quality, or something else in the MLPA reaction.
– Probes are not suitable for non-human DNA.
– Interaction can occur between self designed probes and P200/P300 reference probes.

[4] Signal sloping is the effect that longer probes generate a lower peak height in the electropherogram than shorter probes. The degree of sloping differs between sequencer types. Signal sloping may differ between samples allowing for a bias to occur when reference probes are located only in the longer probe range as is the case in the P200. This is why the P300 reference probemix might be preferred.

[5] To learn more about Q and D control fragments, please consult the MLPA protocol on our website *www.mlpa.com* > Technology > Protocols > MLPA General Protocol (One-Tube).

10.1.1. Adjusting P200/P300 Coffalyser worksheets for inclusion of self-designed probes

Data obtained using the P200 and P300 Reference probemixes with additional synthetic probes can be analysed using Coffalyser.Net, MLPA data analysis software. Coffalyser.Net can be downloaded freely from our website: *http://www.mrcholland.com* (click on Coffalyser.Net). In the Coffalyser Sheet Library you can find the P200 and P300 sheets, to which you can add your synthetic probes by following the steps outlined in the Coffalyser.Net reference manual *Edit Coffalyser sheets*. In addition, it is necessary to manually adjust the role of the control probes in the P200 and P300 Coffalyser sheets as follows: Open the worksheet via the Sheet Library by right-clicking on the sheet, choosing Open and clicking on the probes tab. You will then see the list of probes. In the function(s) column, deselect the "probe" function in the drop down menu for the following probes while leaving the other functions checked for these probes.

- P200: 208 nt X presence control probe, 239 nt Y presence control probe, 251 nt denaturation control probe
- P300: 88 nt denaturation control probe 1, 92 nt ligation control probe, 208 nt X presence control probe, 239 nt Y presence control probe, 252 nt denaturation control probe 2.

Make sure you remember to save the Coffalyser sheet after changing the probe functions and adding additional probes.

10.2. Making an all-synthetic probemix

If you are designing probes for an organism other than human, the only option is to make an all synthetic probemix of self-designed synthetic probes. Fully self-designed synthetic probemixes can be used in combination with an MLPA EK reagent kit. More information on how to pipette this synthetic probemix can be found in section *11 Preparing the synthetic probemix*. This is only recommended for experienced MLPA probe designers.

Advantages (+) & Disadvantages (-) of making an all-synthetic probemix:

+ Suitable for all organisms.
– It is necessary to design reference probes, this can be challenging and takes up design space. It is advised to use a minimum of 8 reference probes (12 for tumour material).
– Difficult to assess whether possible problems are due to synthetic probe design, oligo quality, or other factors.
– No quality control fragments

10.3. EK MLPA reagent kits (for standard MLPA and MS-MLPA)

EK kits contain all necessary MLPA reagents except for a probemix; no reference probes or quality control fragments are present. The MLPA reagents are:

1. MLPA buffer
2. Ligase-65 enzyme
3. Ligase buffer A
4. Ligase buffer B
5. SALSA Polymerase
6. SALSA PCR primers, incl. dNTPs – various fluorescent dyes available, incl FAM, Cy5.0

EK kits are available in the following sizes: EK1 (100 reactions) and EK5 (500 reactions). EK kits can be used both for a standard MLPA reaction as well as Methylation-Specific MLPA (HhaI enzyme can be ordered separately).

11. Preparing the synthetic probemix

11.1. General guidelines

- Do not add more than the recommended amount of each oligo as this may result in non-specific amplification products and lower probe signals. The probe oligo amounts specified below are sufficient to cover >95% of the intended target sequences within the recommended hybridisation time. Having near-complete coverage means that relative probe signals will depend on the relative amount of probe target sequence in the sample and NOT on the exact amount of probe used. A probe signal cannot be increased by the addition of more probe oligos.
- Oligos should be dissolved and diluted in TE (10 mM Tris-HCl, pH=8.0; 1 mM EDTA).
- Dissolved oligos should be stored at -20°C.
- Oligo solutions should never be heated.

11.2. Making a synthetic probemix (step a-c)

Making a synthetic probemix from self-designed oligos requires the following four steps (see **Figure 11** and examples below):

a. Make a 100 µM or 10 µM (ultramers) *oligo stock solution* for each oligo by dissolving oligos in the appropriate amount of TE.
b. Make a 1 µM *working stock solution* for each oligo by diluting the *oligo stock solution* further (either 100x or 10x) with TE.
c. Combine 0.8 µl of each oligo 1 µM *working stock solution* (LPO and RPO of all probes). Once all oligos have been combined add TE to a final volume of 200 µl. This solution is referred to as the *synthetic basic probemix*.
d. Add 0.5 µl of the *synthetic basic probemix* to (1) 1 µl of P200/P300 reference probemix, or (2) dilute it further to use it independently, as specified below. See section *11.3* for more details.

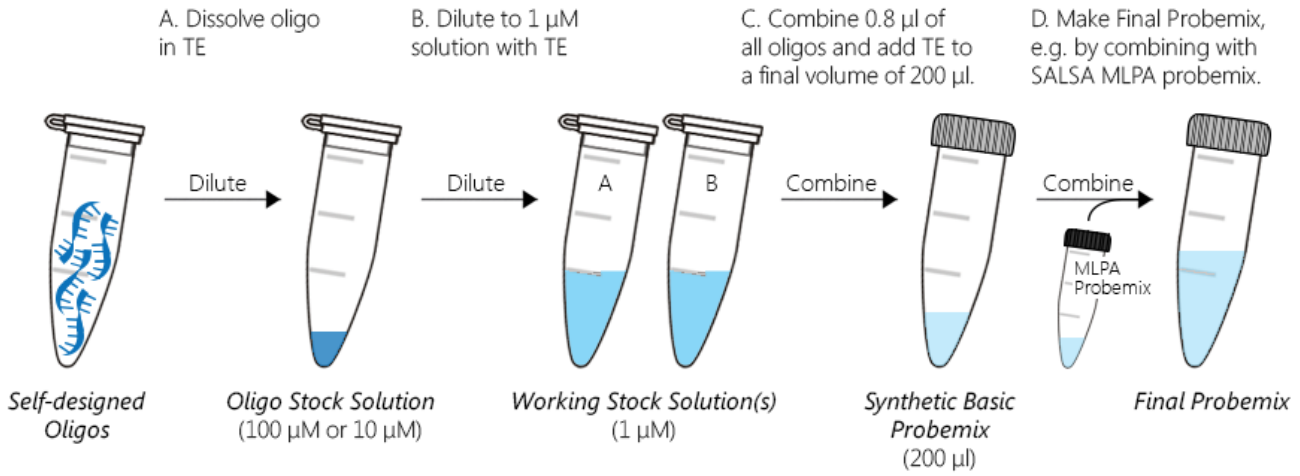


Figure 11 - Workflow of steps for self-designed oligos to the point of inclusion in a probemix that can be used for an MLPA reaction.

Example 1 (step a-c): you receive 40 nmol of lyophilised oligo:

a. Dissolving 40 nmol of oligo in 400 µl TE will result in a 100 µM stock solution.
b. From this 100 µM oligo stock solution, make a 100 fold dilution to a working stock solution (1 µM) of each oligo in TE, for example make 1 ml by diluting 10 µl oligo stock solution with 990 µl TE.
c. Mix 0.8 µl of each 1 µM working stock solution (LPO + RPO), add TE to a total volume of 200 µl.

Example 2 (step a-c): if you receive 4 nmol lyophilised oligo (e.g. IDT ultramers), it is recommended to prepare stock solutions of 10 µM instead of 100 µM.

a. Dissolving 4 nmol of oligo in 400 µl TE will result in a 10 µM stock solution.
b. From this 10 µM oligo stock solution, make a 10 fold dilution to a working stock solution (1 µM) of each oligo in TE, for example make 0.1 ml by diluting 10 µl working stock solution with 90 µl TE.
c. Mix 0.8 µl of each 1 µM working stock solution (LPO + RPO), add TE to a total volume of 200 µl.

11.3. Making the final probemix (step d)

Step d, making the final probemix for use in MLPA reactions depends on what the synthetic oligos are added to, please see the instructions below for the situation that applies to you. Please note that we recommend you make enough probemix for multiple reactions to increase the volumes pipetted.

11.3.1. Situation 1 - Adding oligos to the P200 or P300 MLPA reference probemix:

For each planned MLPA reaction, combine 0.5 µl of your synthetic basic probemix (end product of step c) with 1 µl of P200 or P300 reference probemix. From this solution, use 1.5 µl for each MLPA reaction.

Example 3 - making a synthetic probe mix (step a-d) for one probe (situation 1) with P200/300:

- LPO concentration: received 40.5 nmol, dissolve in 405 µl TE to make a 100 µM *stock solution*.
- RPO concentration: received 27.2 nmol, dissolve in 272 µl TE to make a 100 µM *stock solution*.
- From both the LPO and RPO stock solution, make a 1 µM *working stock solution*. For 1 ml, take 10 µl of the 100 µM *stock solutions* and add to 990 µl TE. Do this for both oligos.
- To obtain a synthetic basic probemix with a total volume of 200 µl, mix:
    - 0.8 µl of the LPO 1 µM working stock solution
    - 0.8 µl of the RPO 1 µM working stock solution
    - 198.4 µl TE
- Step d: use 0.5 µl of this 200 µl synthetic basic probemix together with 1 µl P200 or P300 MLPA reference mix. This combined probemix can be used for an MLPA reaction.

11.3.2. Situation 2 - Making an all-synthetic probemix

Increase the volume of the synthetic basic probemix (end product of step c) to 600* µl by adding TE. From this solution, use 1.5 µl for each MLPA reaction.

All self-designed probemixes should be validated using commercially available DNA to ensure that all probes in the synthetic probemix give a similar probe signal height. If there is a large difference in probe signal height between probes then a competitor oligo mix (COMP) should be used, see section *11.4*.

11.4. Making a competitor oligo mix

A competitor oligo mix can be made and used when the peak height of some/all designed synthetic probes are higher than those of the reference probes, see section *6.6*. It is important that target and reference probes have similar heights for accurate data analysis. The amount of COMP oligo needed in a competitor oligo mix must be adjusted based on the signal intensity of the synthetic probe the COMP is targeting and will likely require optimization. In general, the use of a 1:1 ratio of LPO and its corresponding COMP normally reduces the probe signal approximately 50%. The ratio of each COMP in the competitor oligo mix will likely need to be adjusted after initial use.

a. Dissolve each COMP in TE to a concentration of 100 µM.
b. From this 100 µM COMP stock solution, make a 100 fold dilution working stock solution (1 µM) of each COMP in TE.
c. To achieve a 1:1 ratio of COMP to LPO add 0.8 µl of each 1 µM COMP working stock solution, add TE to a total volume of 200 µl, this is your competitor oligo mix.

d. For each planned MLPA reaction, combine 0.5 µl of your competitor oligo mix, 0.5 µl of your synthetic basic probemix (end product of *11.2* step c), and 1 µl of P200 or P300 reference probemix. From this solution, use 2 µl for each MLPA reaction. Total volume of the hybridisation reaction is increased from 8 to 8.5 µl.

After optimization and determination of the amount of COMP oligo needed per targeted probe it is possible to make a synthetic basic probemix that includes both synthetic probe oligos and COMP oligos. Follow the steps as outlined in *11.2* step a and b then combine 0.8 µl of each oligo *1 µM working stock solution* (LPO and RPO of all probes) and the determined volume of each COMP oligo. Once all oligos have been combined add TE to a final volume of 200 µl. This is now your synthetic basic probemix including COMP oligos and can be used as outlined in section *11.3*.

12. Troubleshooting

All troubleshooting of self-designed synthetic probes/probemixes should be done with commercially available DNA samples.

All probes give low signals:

It is possible all oligos supplied were not of sufficient quality. To test if your oligo supplier provides oligos of good quality, you can order the following probe:

SerpinB2 gene: Chr. 18; length of amplification product: 81+42=123 nt.

LPO: **GGGTTCCCTAAGGGTTGGA**CCATGACTCCAGAGAACTTTACCAGCTGTGGGTTCATGCA

RPO: 5'-P-GCAGATCCAGAAGGGTAGTTATCCTGATGCGATTTTGCAGG**TCTAGATTGGATCTTGCTGGCAC**

The optional competitor sequence that can be used to reduce the signal of this probe (see *6.6 Optional signal-reducing competitor oligo*) is: **TGGA**CCATGACTCCAGAGAACTTTACCAGCTGTGGGTTCATGCA

Low signal of one or a few probes:

If one probe gives a very low signal, addition of up to three times more of that specific probe may sometimes solve the problem, especially for probes targeting GC-rich sequences. A new design of the LPO may also help: remember that having a cytosine as the first nucleotide after the PCR forward primer sequence will generate the highest signal. See *6.3 The effect of the first nucleotide*.

High signal of one or a few probes:

There are two options to reduce peak height. The first nucleotide after the primer of the LPO can be changed to a T, see *6.3 The effect of the first nucleotide*. Another option is to order a competitor oligo, see *6.6 Optional signal-reducing competitor oligo* for more information.

For other MLPA reaction troubleshooting help, see the troubleshooting section on our website.

13. References

The following publications describe the use of synthetic MLPA probes:

- Nicoletti et al. (2017) Design and validation of a new MLPA-based assay for the detection of RS1 gene deletions and application in a large family with X-linked juvenile retinoschisis. *Genet Test Mol Biomarkers* 21(2):116-121.
- Aoyama et al. (2015) Application of multiplex ligation-dependent probe amplification, and identification of a heterozygous Alu-associated deletion and a uniparental disomy of chromosome 1 in two patients with 3-hydroxy-3-methylglutaryl-CoA lyase deficiency. *Int J of Mol Med* 35:1554-1560.

- D'Angelo et al. (2014) Investigation of selected genomic deletions and duplications in a cohort of 338 patients presenting with syndromic obesity by multiplex ligation-dependent probe amplification using synthetic probes. *Mol Cytogenet*. 7:75.
- Kiehntopf et al. (2012) A homemade MLPA assay detects known CTNS mutations and identifies a novel deletion in a previously unresolved cystinosis family. *Gene* 495:89-92.
- Serizawa et al. (2010) Custom-designed MLPA using multiple short synthetic probes: application to methylation analysis of five promoter CpG islands in tumor and urine specimens from patients with bladder cancer. *J Mol Diagn* 12:402-408
- Wildförster, V. and Dekomien, G. (2008) Detecting copy number variations in autosomal recessive limb-girdle muscular dystrophies using a multiplex ligation-dependent probe amplification (MLPA) assay. *Mol Cell Probes*. 23:55-59.
- Roelfsema et al. (2005) Genetic heterogeneity in Rubinstein-Taybi Syndrome: mutations in both the CBP and EP300 genes cause disease. *Am J Hum Genet*. 76:572-580.
- Vink GR et al. (2005) Mutation screening of EXT1 and EXT2 by direct sequence analysis and MLPA in patients with multiple osteochondromas: splice site mutations and exonic deletions account for more than half of the mutations. *Eur J Hum Genet*. 13:470-474.
- Langerak P et al. (2005) Rapid and quantitative detection of homologous and non-homologous recombination events using three oligonucleotide MLPA. *Nucleic Acids Res*. 33:e188.
- Stern RF et al. (2004) Multiplex ligation-dependent probe amplification using a completely synthetic probe set. *Biotechniques*. 37:399-405.

14. Useful websites, tools and software

If a URL is no longer working, please notify us at *info@mrcholland.com*.

Table 4- Programs used for synthetic probe design

Coffalyser.Net	Free software for analysis of MLPA results, available via *www.mrcholland.com* → Coffalyser.Net
Mfold DNA Folding Program	*http://mfold.rna.albany.edu/?q=mfold* Parameters should be set on [Na+] = 0.35 M, T = 60° C.
Ensembl	*http://www.ensembl.org/index.html* Ensembl is a joint project between EMBL - EBI and the Sanger Institute. It is helpful for information on certain genes
Entrez Gene	*https://www.ncbi.nlm.nih.gov/gene*, useful for finding NM-sequences (for exon numbers, see NG-sequences: *http://www.ncbi.nlm.nih.gov/refseq/rsg/browse/*).
GenBank	*https://www.ncbi.nlm.nih.gov/nucleotideUseful* for looking up reference (NM_) sequences.
Genomic Variants Database	*http://dgv.tcag.ca/dgv/app/home* Database of copy number variants in the human genome
Genome Data Viewer	*https://www.ncbi.nlm.nih.gov/genome/gdv/*
Human Genome BLAST	*http://www.ncbi.nlm.nih.gov/genome/seq/BlastGen/BlastGen.cgi?taxid=9606*
Human Genome Organization	*http://www.genenames.org*
Integrated DNA Technologies	*www.idtdna.com*
Locus Reference Genomic	*https://www.lrg-sequence.org/index.html*
NR-BLAST	*https://blast.ncbi.nlm.nih.gov/Blast.cgi*
OMIM	*https://www.ncbi.nlm.nih.gov/omim*
PubMed	*https://www.ncbi.nlm.nih.gov/pubmed*
RaW-Probe	*http://www.support.mlpa.com* – Downloads – Synthetic Probe Design Download RaW-Probe program
UCSC BLAT	*http://genome.ucsc.edu/cgi-bin/hgBlat?command=start*

Please note that designing MLPA for commercial use is not allowed when this infringes on MLPA patents US 6955901, CA2400240, EP 1130 113 A1, US 2007009288.

Implemented Changes compared to previous synthetic probe design protocol versions
Version 4 – November 2018 - Company website address changed to *www.mrcholland.com*. - Company email address changed to *info@mrcholland.com*. - Removed information about CF1 and CF4 control fragment mixes, because these are discontinued. - Minor formatting changes. *Version 3 – September 2018* - Number of reference probes and total probes in a synthetic mix revised - reducing mismatch signal revised *Version 2 – 13 July 2018* - New document template applied - Figures and images updated - P200/300 tables removed - Numerous textual revisions made - links updated *Version 1 (equal to version 15 in previous template) - June 2017* - New document template applied.