Increasingly, classification of prokaryotes has involved molecular phylogenetic methods. This identification invariably relies on a genotypic approach, typically involving an analysis of 16s rRNA gene. More frequently, 16S rRNA information is used to identify oligonucleotide sequences unique to specific bacteria, for use as hybridization probes or as PCR primers. Such oligonucleotides can be specific to phylogenetic groupings as diverse as bacterial species or an entire Phylum. Furthermore, they are regularly used to enumerate specific groups of bacteria in natural environments by fluorescence in situ hybridization (FISH) technique and also in the gene expression studies of several genome projects. In this regard, a good probe design plays a crucial role in determining a potential probe that matches as many of the intended target group as possible, while ignoring the non-members of the target group.

ARB software package provides a workbench, where probes can be designed, evaluated and visualized in more intuitive way, using interactive graphical user interface (GUI) of ARB. Algorithms of the ARB programs ‘Probe Design’ and ‘Probe Match’ are searching the PT-server to identify short (10 - 100 monomers) diagnostic sequence stretches which are evaluated against the background of all full and partial sequences in the respective database. The PT-server has been built from the principle, no alignment of the sequence data is needed for specific probe design. However, in case of taxonomic specific probes, alignment and phylogenetic analyses are necessary to allow defining groups of phylogenetically (taxonomically) related organisms as the targets of specific probes. The design of taxonomic specific oligonucleotide probes with ARB is performed in three steps. Firstly, the user selects the organism or a group of organisms for which he wants to design a diagnostic probe. Secondly, the software ‘Probe Design’ searches the PT-server for potential target sites. The results are shown in a ranked list of proposed targets, probes and additional information. The ranking is according to several compositional and thermodynamical criteria. Thirdly, the proposed oligonucleotide probes are evaluated against the whole database by using the program “Probe Match”. Local alignments are determined between the probe target sequences and the most similar reference sequences (optionally from 0 to 5 mismatches) in the respective database. Furthermore, these sequence strings can automatically be visualized in the primary and secondary structure editors. Especially the latter information is of high importance when designing probes for in situ cell hybridisation. A special advancement is the ARB multi-probe software component. It determines sets of up to five probes optimally identifying the target group. This probe set can be used for multiple fluorescence in situ hybridisation experiments.

As an example, probes were designed for the Enterobacteria group (colored in fluorescent green) in the phylogenetic tree (16S Dendrogram tree view) of small subunit (16S) ribosomal RNA. The ARB Project is based on the concept of integrated databases of raw and/or processed sequence and any other type of additional data assigned to the individual sequence entries. The workbench facilitates curation, alignment, visualization, and analysis of databases of any genes or genomes. This is provided by a selection of evolutionary models (secondary structure of ribosomal RNA, consensus models), columns (maximum column base frequency, Positional variability model) and any other user defined models, filters or statistics, in different colors. Additionally, user has an option to visualize the sequence associated information below each target species. In this screenshot, well accepted secondary structure model of 16S rRNA is visualized in red-color (helix region), where the starting and ending positions of helices are colored in black, and bases without background represents the loop regions of the structure.

The ARB (derived from the Latin word, arbor, meaning tree) Project, was started as an interdisciplinary initiative of the Lehrstuhl für Mikrobiologie and the Lehrstuhl für Rechnerotechnik und Rechnerorganisation, at the Technical University of Munich, more than a decade ago. The ARB software package provides with an extensive and greedy software environment for the researchers in the area of Molecular Phylogeny, Molecular Ecology and Molecular Taxonomy. The ARB Project is based on the concept of integrated databases of raw and/or processed sequence and any other type of additional data assigned to the individual sequence entries. The workbench facilitates curation, alignment, visualization, and analysis of databases of any genes or genomes. This is provided by a selection of evolutionary models (secondary structure of ribosomal RNA, consensus models), columns (maximum column base frequency, Positional variability model) and any other user defined models, filters or statistics, in different colors. Additionally, user has an option to visualize the sequence associated information below each target species.