	EC numbers	Swiss-Prot IDs	Aligned positions corresponding to Functionally Discriminating Residues (FDRs)					
			113	124	218	238	240	302
Homofunctional Multiple Sequence Alignment	1.1.1.28	LDHD_ECOLI	PEAVA**E**HAIGM	MMTLN**R**RIHRA	PLTPE**N**YHLLN	GVMIV**N**TS**R**GA		VLFTG**H**QAFLT
	1.1.1.28	LDHD_LACDE	PNAIA**E**HAAIQ	AARIL**R**QAKAM	PDVPA**N**VHMIN	DVVIV**N**VS**R**GP		VLVTP**H**TAFYT
	1.1.1.28	LDHD_LACHE	PNAIA**E**HAAIQ	AARVL**R**QDKRM	PDVPA**N**VHMIN	GVVIV**N**CS**R**GR		VLVTP**H**TAFYT
	1.1.1.28	LDHD_LACPE	PNAIA**E**LSVTQ	LMQLL**R**QTPMF	PALKD**N**YHMLN	GAYIL**N**FA**R**GT		VLITP**H**TAFYT
	1.1.1.28	LDHD_LEUMC	PNAIA**E**HSMLQ	LSRLL**R**RTKAL	PGVPE**N**HHLIN	GVVIM**N**AA**R**GN		VMITP**H**TAFYT
	1.1.1.28	LDHD_PEDAC	PRAIA**E**LSVTQ	LLALL**R**KIPEF	PALKD**N**YHMLD	GTFIL**N**FA**R**GT		VMITP**H**AAFYT
Heterofunctional Multiple Sequence Alignment	1.1.1.95	SERA_METTH	SITVA**E**HSIGL	MLALA**R**KIAIA	PLTPE**T**RHLIS	TAFIV**N**CA**R**GG		VVLTP**H**IGAST
	3.3.1.1	SAHH_SYNY3	-RAMF**N**DGVLT	FPAMN**V**NDHFY	--VTG**N**KHVIR	GAIVC**N**SG**H**--		----G**H**-----
	1.1.1.28	LDHD_ECOLI	PEAVA**E**HAIGM	MMTLN**R**RIHRA	PLTPE**N**YHLLN	GVMIV**N**TS**R**GA		VLFTG**H**QAFLT
	1.20.1.1	PTXD_PSEST	TVPTA**E**LAIGL	AVGLG**R**HLRAA	PLNAD**T**QHLVN	GALLV**N**PC**R**GS		TLFTP**H**IGSAV
	1.2.1.2	FDH_EMENI	VVSVA**E**HVVMT	ILLLV**R**NFVPA	PLHEK**T**RGLFN	GSWLV**N**TA**R**GA		-----**H**PWGGG
	1.2.1.2	FDH_PICAN	SNSVA**E**HVVMT	MLVLV**R**NFVPA	PLHAG**S**KGLVN	GAWLV**N**TA**R**GA		AM-TP**H**Y----
	1.1.1.29	DHGY_METEX	FNTVP**E**HVVGL	MFALR**R**AIVPY	PLTPD**T**KNMIG	SAILI**N**TA**R**GG		--VTP**H**VAWAS
	1.1.1.215	TKRA_BACSU	DDTVA**D**LAFSL	ILSSA**R**RVAEL	PLTDE**T**YHMIG	SAIFV**N**IS**R**GK		VTLLP**H**IGSAT
	non-enz.	CTBP_XENLA	VEETA**D**STMCH	ILNLY**R**RTTWL	GLNEH**N**HHLIN	GAFLV**N**TA**R**GG		LICTP**H**AAWYS
	....							

Figure S1. Example of Functionally Discriminating Residues (FDRs). Aligned positions (numbered boxes outlined in red) corresponding to FDRs in the Homofunctional Multiple Sequence Alignment (MSA) and Heterofunctional MSA for an enzyme family whose seed EC number is 1.1.1.28 (D-lactate dehydrogenase). The six FDRs are selected for their ability to discriminate between Homofunctional (i.e., D-lactate dehydrogenases) and Heterofunctional family members (i.e., evolutionary related proteins with any other function). We assign EC number 1.1.1.28 to any sequence recognized as member of the family by a Hidden Markov Model and showing conservation of the six FDRs: E113, R124, N218, N238, R240 (which corresponds to an active site residue in D-lactate dehydrogenases, e.g. active site residue R234 in LDHD_ECOLI) and H302.