Cooperative Binding of PhoB^{DBD} to its Cognate DNA Sequence - A Combined Application of Single-Molecule and Ensemble Methods.

Markus Ritzefeld[†], Volker Walhorn[‡], Christin Kleineberg[‡], Adeline Bieker[‡], Klaus Kock[§], Christian Herrmann[§], Dario Anselmetti[‡] and Norbert Sewald[†].*

[†] Organic and Bioorganic Chemistry, Bielefeld University, Universitaetsstr. 25, 33615 Bielefeld, Germany

[‡] Experimental Biophysics, Bielefeld University, Universitaetsstr. 25, 33615 Bielefeld, Germany

[§] Physical Chemistry I, Ruhr University Bochum, Universitaetsstr. 150, 44801 Bochum, Germany

[*] Prof. Dr. Norbert Sewald, Organic and Bioorganic Chemistry, Bielefeld University, Universitaetsstr. 25, 33615 Bielefeld, Germany. Tel.: +49 521 106 2051, E mail: Norbert.Sewald@uni-bielefeld.de

Supporting Information

S1: MALDI-ToF MS Results

Protein	Calculated [m/z]	Measured [m/z]
Q135C	11999	12003
PhoB-Atto550	12717	12729
PhoB-CF	12673	12673
R176A	11939	11937
T217A	11994	11992
V218A	11996	11996
Y223A	11932	11930
CRA-PhoB^{DBD}	12486	12489
CRA-LEAF-PhoB^{DBD}	12946	12947
CRA-PhoB^{DBD}-dimer	24549	24563
CRA-LEAF-PhoB^{DBD}-dimer	25009	24993

S2: Results of the Tryptic Digestion of PhoB^{DBD}-dimers

CRA-PhoB^{DBD}-dimer

AVEEVIEMQG LSLDPTSHRV MAGEEPLEMG PTEFKLLHFF MTHPERVYSR EQLLNHVWGT
NVYVEDRTVD VHIRRLRKAL EPGGHDRMVQ TVRGTGYRFS TRFC*RAMAVE EVIEMQGLSL
DPTSHRVMAG EEPLEMGPTE FKLLHFFMTH PERVYSREQL LNHVWGTNVY VEDRTVDVHI
RRLRKALEPG GHDRMVQTVR GTGYRFSTRF

* Carbamidomethyl modification

Sequence Coverage: 96.7%

The peptide fragment containing the ligation site is highlighted in bold.

Measured [m/z]	Calculated [m/z]	Δ [ppm]	Rel. Int. [%]	Annotation	Formula
510.125	510.2671	-278.5	4.27	[206-209] r.FSTR.f	$C_{22}H_{35}N_7O_7$
510.125	510.2671	-278.5	4.27	[99-102] r.FSTR.f	$C_{22}H_{35}N_7O_7$
524.2944	524.2827	22.4	32.09	[47-50] r.VYSR.e	$C_{23}H_{37}N_7O_7$
524.2944	524.2827	22.4	32.09	[154-157] r.VYSR.e	$C_{23}H_{37}N_7O_7$
553.2787	553.2729	10.5	24.87	[94-98] r.GTGYR.f	$C_{23}H_{36}N_8O_8$
553.2787	553.2729	10.5	24.87	[201-205] r.GTGYR.f	$C_{23}H_{36}N_8O_8$
733.3879	733.4025	-20	58.67	[88-93] r.MVQTVR.g	$C_{30}H_{56}N_{10}O_9S$
733.3879	733.4025	-20	58.67	[195-200] r.MVQTVR.g	$C_{30}H_{56}N_{10}O_9S$
839.4558	839.4734	-20.9	100	[68-74] r.TVDVHIR.r	$C_{36}H_{62}N_{12}O_{11}$
839.4558	839.4734	-20.9	100	[175-181] r.TVDVHIR.r	$C_{36}H_{62}N_{12}O_{11}$
951.4373	951.4643	-28.3	13.9	[79-87] k.ALEPGGHDR.m	$C_{39}H_{62}N_{14}O_{14}$
951.4373	951.4643	-28.3	13.9	[186-194] k.ALEPGGHDR.m	$C_{39}H_{62}N_{14}O_{14}$
973.4184	**973.4672**	**-50.2**	**3.83**	**[99-105] r.FSTRFC*R.a [*Carbamidomethyl]**	**$C_{42}H_{64}N_{14}O_{11}S$**
1079.5311	1079.5592	-26	3.64	[78-87] r.KALEPGGHDR.m	$C_{45}H_{74}N_{16}O_{15}$
1079.5311	1079.5592	-26	3.64	[185-194] r.KALEPGGHDR.m	$C_{45}H_{74}N_{16}O_{15}$
1427.6939	1427.7253	-21.9	97.39	[36-46] k.LLHFFMTHPER.v	$C_{67}H_{98}N_{18}O_{15}S$
1427.6939	1427.7253	-21.9	97.39	[143-153] k.LLHFFMTHPER.v	$C_{67}H_{98}N_{18}O_{15}S$
1764.7769	1764.8183	-23.5	0.74	[20-35] r.VMAGEEPLEMGPTEFK.l	$C_{77}H_{121}N_{17}O_{26}S_2$
1764.7769	1764.8183	-23.5	0.74	[127-142] r.VMAGEEPLEMGPTEFK.l	$C_{77}H_{121}N_{17}O_{26}S_2$
2071.9586	2072.0196	-29.4	20.52	[51-67] r.EQLLNHVWGTNVYVEDR.t	$C_{92}H_{138}N_{26}O_{29}$
2071.9586	2072.0196	-29.4	20.52	[158-174] r.EQLLNHVWGTNVYVEDR.t	$C_{92}H_{138}N_{26}O_{29}$
2110.9565	2111.0437	-41.3	23.99	[1-19] .AVEEVIEMQGLSLDPTSHR.v	$C_{89}H_{147}N_{25}O_{32}S$
2313.0545	2313.1213	-28.9	16.86	[106-126] r.AMAVEEVIEMQGLSLDPTSHR.v	$C_{97}H_{161}N_{27}O_{34}S_2$

CRA-LEAF-PhoB^{DBD}-dimer

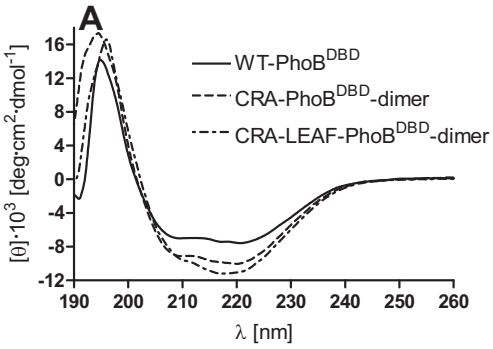
AVEEVIEMQG LSLDPTSHRV MAGEEPLEMG PTEFKLLHFF MTHPERVYSR EQLLNHVWGT
NVYVEDRTVD VHIRRLRKAL EPGGHDRMVQ TVRGTGYRFS TRFC*RALEAF MAVEEVIEMQ
GLSLDPTSHR VMAGEEPLEM GPTEFKLLHF FMTHPERVYS REQLLNHVWG TNVYVEDRTV
DVHIRRLRKA LEPGGHDRMV QTVRGTGYRF STRF

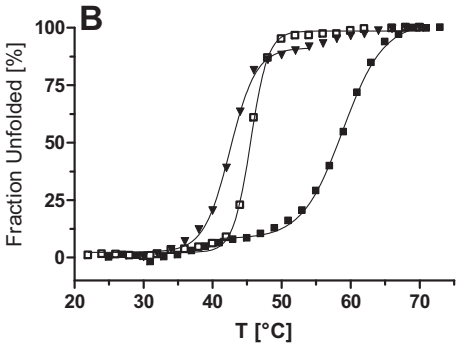
* Carbamidomethyl modification

Sequence Coverage: 97.2%

The peptide fragment containing the ligation site is highlighted in bold.

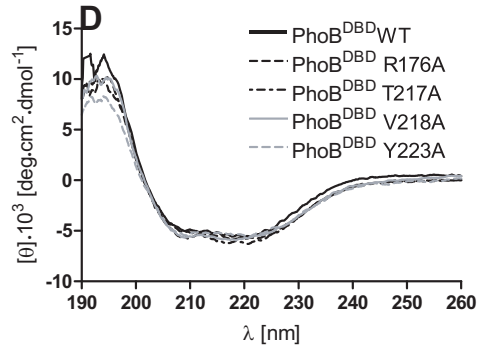
Measured [m/z]	Calculated [m/z]	Δ [ppm]	Rel. Int. [%]	Annotation	Formula
510.1469	510.2671	-235.5	23.96	[99-102] r.FSTR.f	$C_{22}H_{35}N_7O_7$
510.1469	510.2671	-235.5	23.96	[210-213] r.FSTR.f	$C_{22}H_{35}N_7O_7$
524.1674	524.2827	-220	27.57	[158-161] r.VYSR.e	$C_{23}H_{37}N_7O_7$
524.1674	524.2827	-220	27.57	[47-50] r.VYSR.e	$C_{23}H_{37}N_7O_7$
552.8942	553.2729	-684.4	4.07	[205-209] r.GTGYR.f	$C_{23}H_{36}N_8O_8$
552.8942	553.2729	-684.4	4.07	[94-98] r.GTGYR.f	$C_{23}H_{36}N_8O_8$
553.1687	553.2729	-188.4	29.43	[94-98] r.GTGYR.f	$C_{23}H_{36}N_8O_8$
553.1687	553.2729	-188.4	29.43	[205-209] r.GTGYR.f	$C_{23}H_{36}N_8O_8$
657.2605	657.3355	-114.1	4.04	[210-214] r.FSTRF.	$C_{31}H_{44}N_8O_8$
732.9306	733.4025	-643.5	4.04	[88-93] r.MVQTVR.g	$C_{30}H_{56}N_{10}O_9S$
732.9306	733.4025	-643.5	4.04	[199-204] r.MVQTVR.g	$C_{30}H_{56}N_{10}O_9S$
733.3451	733.4025	-78.3	25.56	[199-204] r.MVQTVR.g	$C_{30}H_{56}N_{10}O_9S$
733.3451	733.4025	-78.3	25.56	[88-93] r.MVQTVR.g	$C_{30}H_{56}N_{10}O_9S$
839.4377	839.4734	-42.5	83.37	[68-74] r.TVDVHIR.r	$C_{36}H_{62}N_{12}O_{11}$
839.4377	839.4734	-42.5	83.37	[179-185] r.TVDVHIR.r	$C_{36}H_{62}N_{12}O_{11}$
951.4332	951.4643	-32.7	9.79	[79-87] k.ALEPGGHDR.m	$C_{39}H_{62}N_{14}O_{14}$
951.4332	951.4643	-32.7	9.79	[190-198] k.ALEPGGHDR.m	$C_{39}H_{62}N_{14}O_{14}$
973.4149	**973.4672**	**-53.7**	**3.37**	**[99-105] r.FSTRFC*R.a [*Carbamidomethyl]**	**$C_{42}H_{64}N_{14}O_{11}S$**
1079.5328	1079.5592	-24.5	4.21	[78-87] r.KALEPGGHDR.m	$C_{45}H_{74}N_{16}O_{15}$
1079.5328	1079.5592	-24.5	4.21	[189-198] r.KALEPGGHDR.m	$C_{45}H_{74}N_{16}O_{15}$
1427.715	1427.7253	-7.2	100	[147-157] k.LLHFFMTHPER.v	$C_{67}H_{98}N_{18}O_{15}S$
1427.715	1427.7253	-7.2	100	[36-46] k.LLHFFMTHPER.v	$C_{67}H_{98}N_{18}O_{15}S$
1764.8005	1764.8183	-10.1	3.97	[131-146] r.VMAGEEPLEMGPTEFK.l	$C_{77}H_{121}N_{17}O_{26}S_2$
1764.8005	1764.8183	-10.1	3.97	[20-35] r.VMAGEEPLEMGPTEFK.l	$C_{77}H_{121}N_{17}O_{26}S_2$
2072.0002	2072.0196	-9.4	65.08	[51-67] r.EQLLNHVWGTNVYVEDR.t	$C_{92}H_{138}N_{26}O_{29}$
2072.0002	2072.0196	-9.4	65.08	[162-178] r.EQLLNHVWGTNVYVEDR.t	$C_{92}H_{138}N_{26}O_{29}$
2111.0057	2111.0437	-18	47.51	[1-19] .AVEEVIEMQGLSLDPTSHR.v	$C_{89}H_{147}N_{25}O_{32}S$
2773.2837	2773.3535	-25.2	3.54	[106-130]r.ALEAFMAVEEVIEMQGLSLDPTSHR.v	$C_{120}H_{193}N_{31}O_{40}S_2$





C

Protein	Helix [%]	β-sheet [%]	Turn[%]	Random [%]		T [°C]	CI [°C]	R^2
WT-PhoB^{DBD}	31.9	34.7	12.2	21.2	■	58.8	58.5-59.1	99.9
CRA-PhoB^{DBD}-dimer	32.2	38.8	4.9	23.9	□	45.5	45.3-45.6	99.9
CRA-LEAF-PhoB^{DBD}-dimer	26.2	35.3	18.3	20.2	▼	42.4	42.2-42.8	99.8



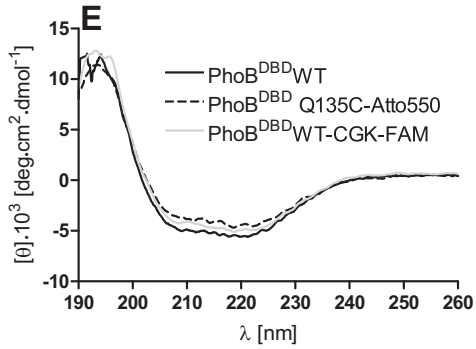


Figure S3. CD Spectra of all investigated proteins. **A** Overlay of the CD spectra of WT-PhoB^{DBD} and the two dimers CRA-PhoB^{DBD} and CRA-LEAF-PhoB^{DBD}. **B** The course of the CD effect at 220 nm was normalized to the fraction of unfolded protein. **C** The fraction of all secondary structure elements was calculated from the CD spectra using Yang's reference [21]. The temperature of heat denaturation was elucidated from the slope of the fraction unfolded temperature plot using nonlinear regression (CI = 95% confidence interval). **D** CD spectra of WT-PhoB^{DBD} and the corresponding point mutants R176A, T217A, V218A and Y223A. **E** CD spectra of WT-PhoB^{DBD} and the two fluorescently labeled proteins (PhoB-CF and PhoB-Atto550).

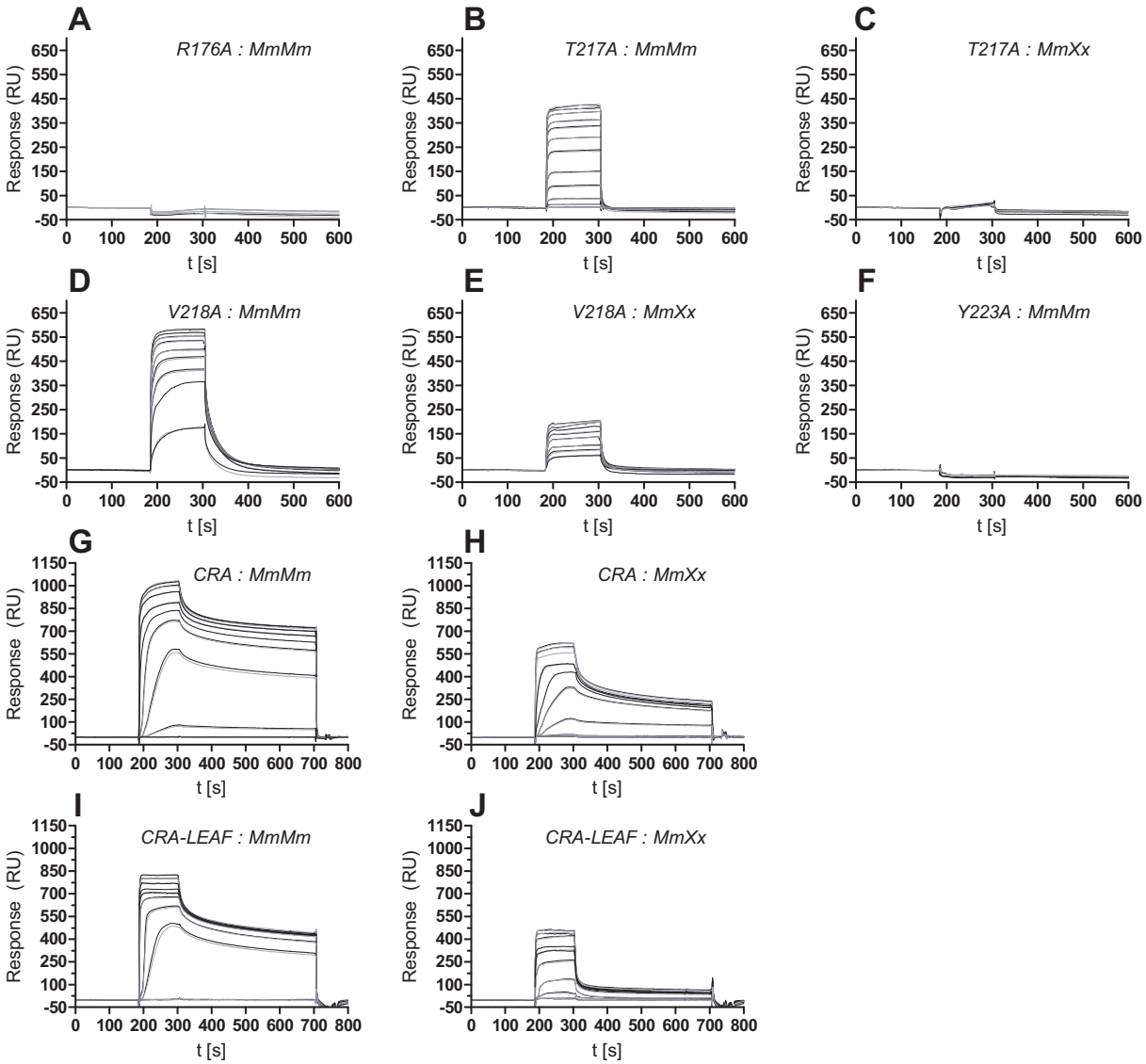


Figure S4. Sensograms of the interaction between PhoB^{DBD} epitopes and the oligonucleotides *MmMm* and *MmXx*. The proteins R176A (**A**), T217A (**B, C**), V218A (**D, E**), Y223A (**F**), CRA-PhoB^{DBD}-dimer (**G, H**) and CRA-LEAF-PhoB^{DBD}-dimer (**I, J**) were used as analytes at different concentrations between 100 and 0.5 µM in the case of the point mutants and 10 and 0.05 µM in the case of the dimers, respectively. All sensograms display both runs of every analyte concentration colored in black and grey, to show the reproducibility of the method.

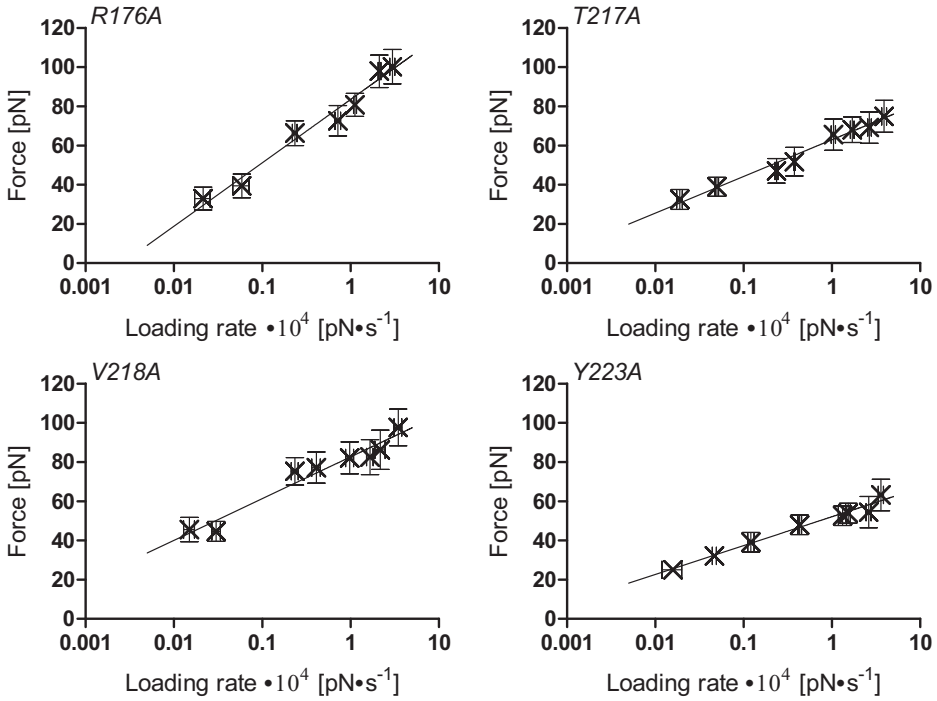


Figure S5. Half-logarithmic DFS force versus loading rate plot of the protein-DNA complexes consisting of the different point mutants and the oligonucleotide (268 bp *pstS pho* box).

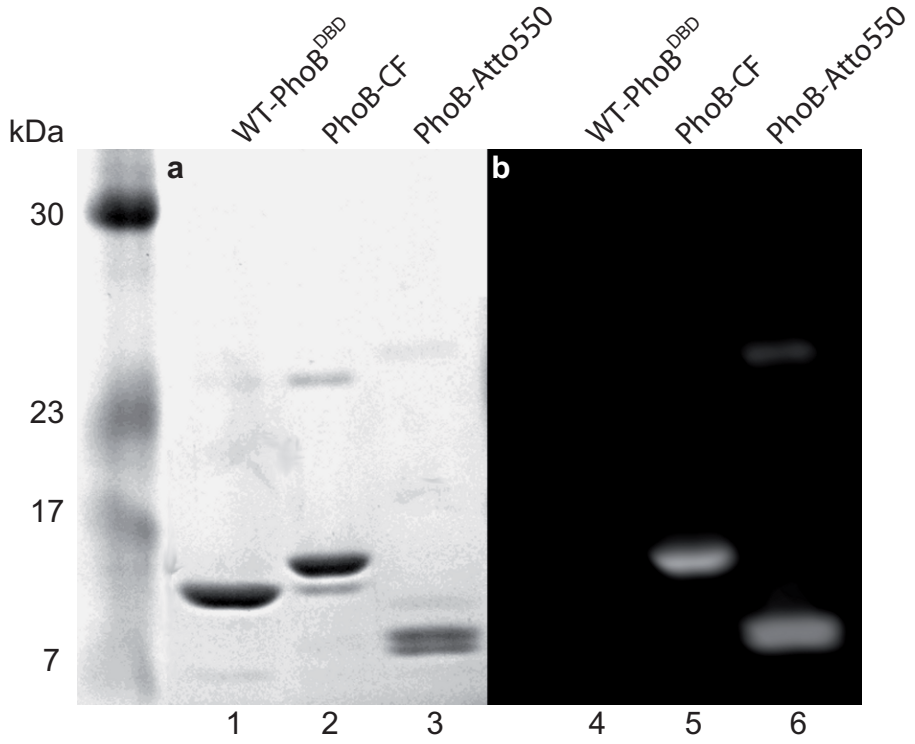


Figure S6. 10% SDS-PAGE. The first three lanes of the SDS-PAGE show the coomassie-stained gel and the last three lanes an overlay of two fluorescence pictures. **Lane 1** and **4** exhibit the WT-PhoB^{DBD} protein (12 kDa) for comparison. The CF-labeled protein is clearly visible in the fluorescence picture in **lane 5** (Excitation: 488 nm, Emissionfilter: 520 nm; 12.7 kDa). The corresponding coomassie-stained gel shows two very weak impurities in lane 2 that correspond to a marginal amount of unlabeled WT-PhoB^{DBD} (12 kDa) and the chitin binding domain plus intein (25 kDa). Traces of the latter always remain after the purification by chitin affinity chromatography.[18] **Lane 6** shows the fluorescence picture of the Atto 550 labeled protein (Excitation: 532 nm, Emissionfilter: 580 nm). The traces of the chitin binding domain plus intein (25 kDa) were also labeled by the Atto dye due to the existence of a cysteine residue. Interestingly, only two other bands are visible in **lane 3** and **6** at approximately 10 kDa. Although they exhibit a molecular weight divergence with respect to the target protein of approximately 4 kDa, tryptic digestion followed by mass spectrometry analysis of both bands revealed with a PhoB^{DBD} sequence coverage and MASCOT score of 79% and 139, respectively, that both bands correspond to the labeled protein. One possible explanation for the deviant running characteristics is an inhomogeneous coverage of the protein with SDS due to the Atto dye.

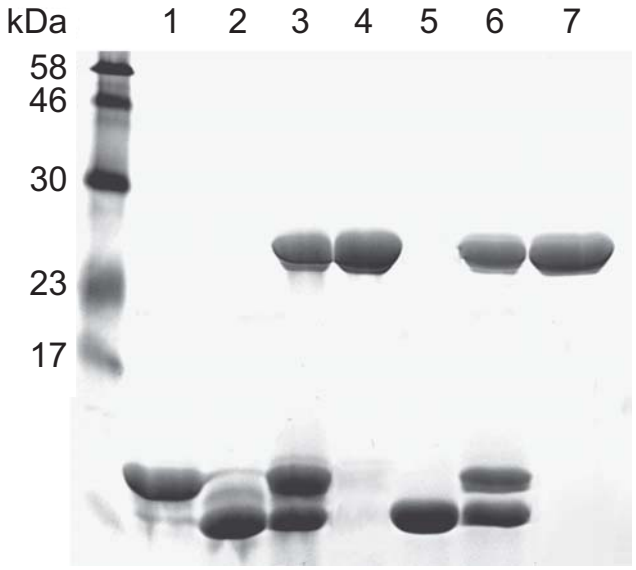


Figure S7. SDS-PAGE (15%) of the two dimer preparations. **Lane 1**: WT-PhoB^{DBD}, **lane 2**: CRA-PhoB^{DBD}, **lane 3**: reaction mixture of the preparation of CRA-PhoB^{DBD}-dimer, **lane 4**: CRA-PhoB^{DBD}-dimer after purification, **lane 5**: CRA-LEAF-PhoB^{DBD}, **lane 6**: reaction mixture of the preparation of CRA-LEAF-PhoB^{DBD}-dimer, **lane 7**: CRA-LEAF-PhoB^{DBD}-dimer after purification. Although WT-PhoB^{DBD} exhibits a smaller molecular weight in comparison to CRA-PhoB^{DBD} and CRA-LEAF-PhoB^{DBD}, the SDS-PAGE indicates the opposite (cf. lanes 1, 2 and 5). Peptide mass fingerprinting analysis after tryptic digestion was performed to verify the data. The bands of the two cysteine epitopes correspond to the correct protein sequences (data not shown). In conclusion, the observed gel shift might be due to an irregularly binding of the detergent SDS, a well-known problem e.g. in the case of membrane proteins.