# Structural Basis of Gene Regulation by the Grainyhead/CP2 Transcription Factor Family

Qianqian Ming[1,2], Yvette Roske[1], Anja Schuetz[1,3], Katharina Walentin[4], Ibraim Ibraimi[4], Kai M. Schmidt-Ott[4,5] and Udo Heinemann[1,2,3,*]

[1] Macromolecular Structure and Interaction, Max Delbrück Center for Molecular Medicine, Robert-Rössle-Str. 10, 13125 Berlin, Germany
[2] Chemistry and Biochemistry Institute, Freie Universität Berlin, Takustr. 6, 14195 Berlin, Germany
[3] Helmholtz Protein Sample Production Facility, Max Delbrück Center for Molecular Medicine, Robert-Rössle-Str. 10, 13125 Berlin, Germany
[4] Molecular and Translational Kidney Research, Max Delbrück Center for Molecular Medicine, Robert-Rössle-Str. 10, 13125 Berlin, Germany
[5] Department of Nephrology, Charité Medical University, Charitéplatz 1, 10117 Berlin, Germany

* To whom correspondence should be addressed. Tel: +49 30 9406 3420; Fax: +49 30 9406 2548
Email: heinemann@mdc-berlin.de

## Supporting Information Appendix

**SUPPLEMENTARY FIGURES AND TABLES**

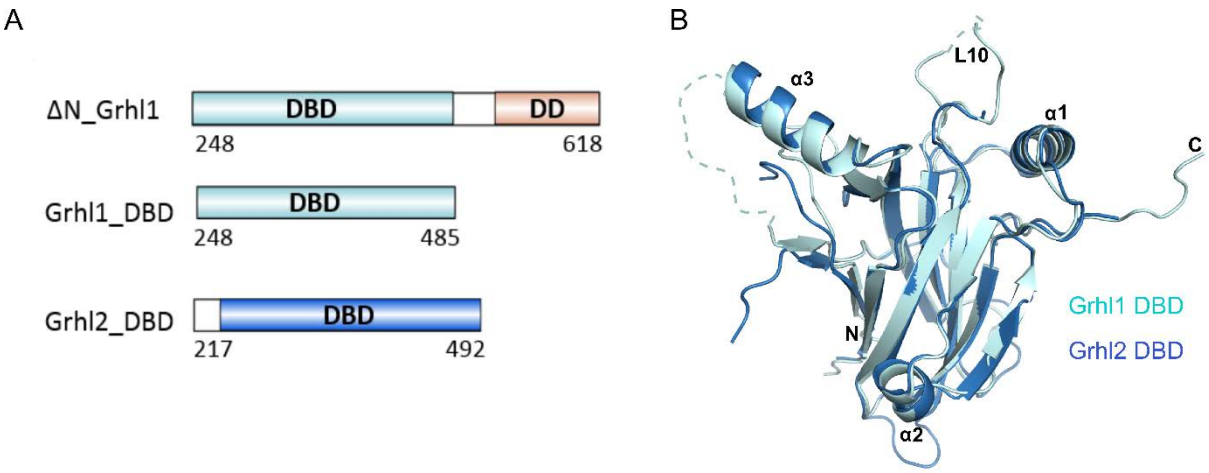



**Fig. S1.** The DNA-binding domain is structurally conserved in Grhl1 and Grhl2. (*A*) Recombinant constructs of Grhl1 and Grhl2 used for crystallization and biochemical studies. (*B*) Superimposition of Grhl1 (pale cyan) and Grhl2 (chain B, blue) DBD reveals a closely similar protein structure.

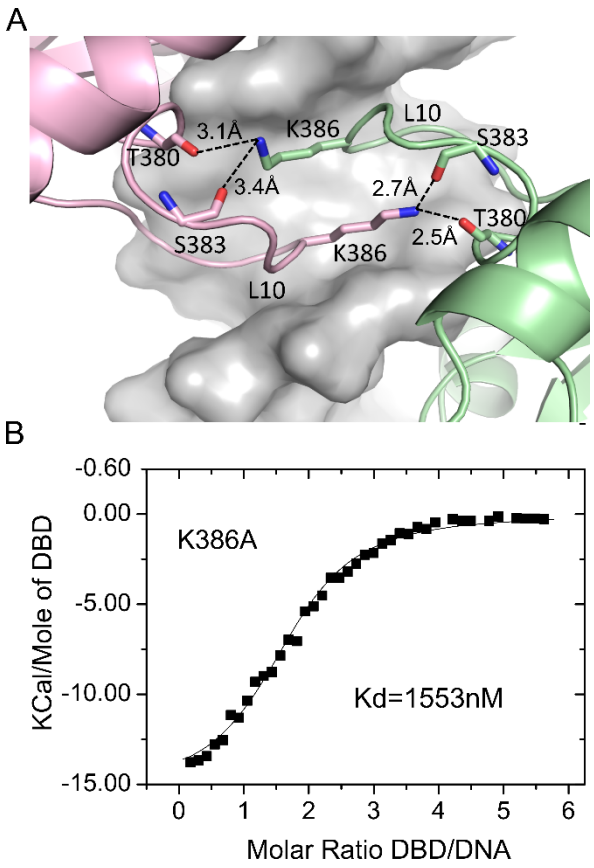



**Fig. S2.** Disruption of protein-protein interface results in dramatic reduction in DNA-binding affinity of Grhl1 DBD. (*A*) Protein-protein interactions found in the Grhl1-DBD:DNA complex structure. Lys386 within the L10 loop of Grhl1 DBD forms hydrogen bonds with Thr380 and Ser383 from the opposite Grhl1-DBD molecule. The DNA surface is shown in gray. (*B*) ITC measurement of the DNA binding affinity of mutant Grhl1-DBD K386A. The Kd value is increased by 17 fold compared to wildtype Grhl1 DBD.

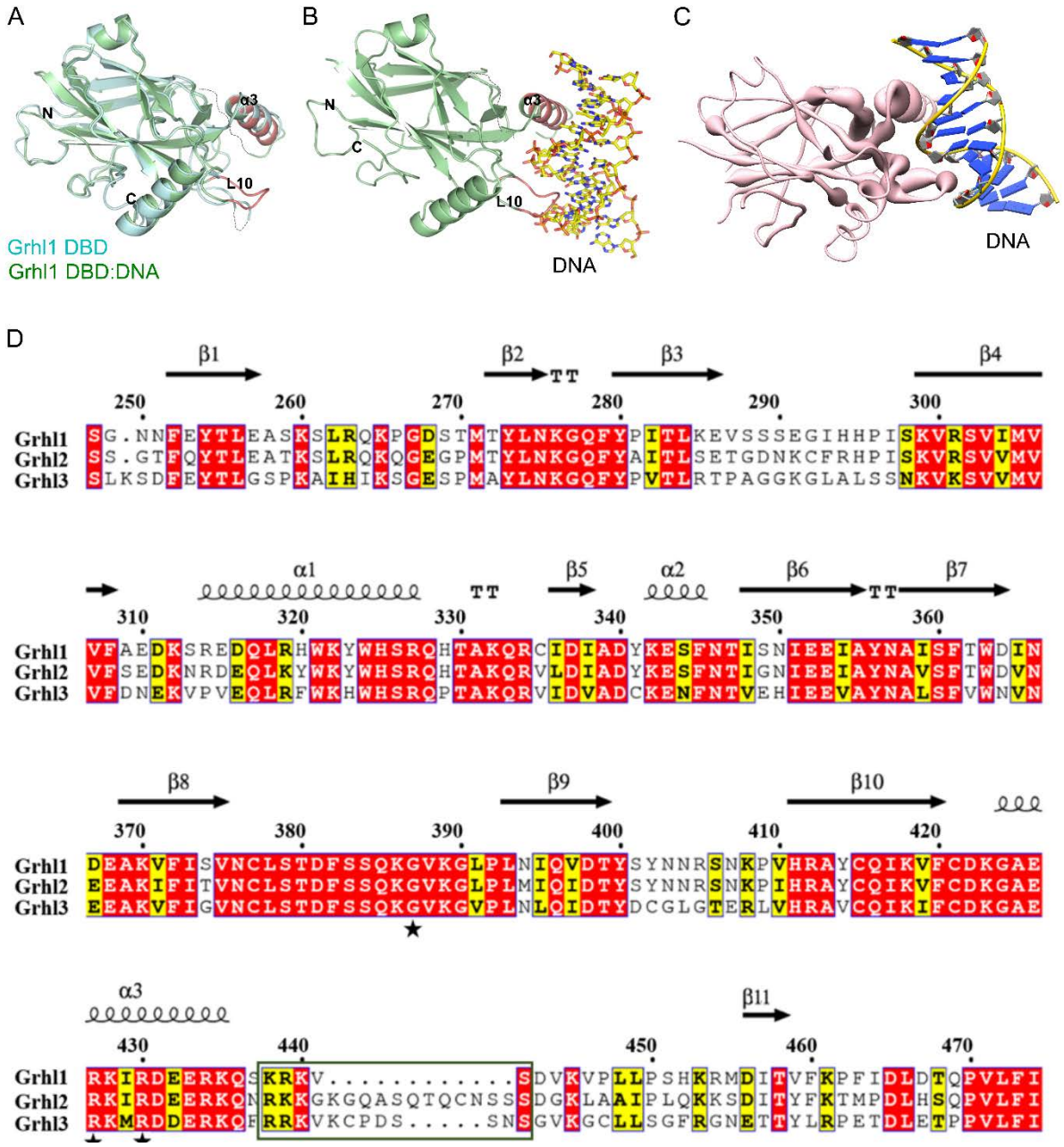



**Fig. S3.** Grhl1 DBD adopts a preformed conformation and conserved interface for DNA binding. (*A*) Structural comparison of free and DNA-bound Grhl1 DBD, showing a mean RMSD for matching Cα atoms of 0.56 Å and a marginally changed orientation of DNA-contacting helix α3 and loop L10. (*B*) Loop L10, contacting DNA in the minor groove, becomes less flexible upon DNA binding. (*C*) Tube representation of Grhl1-DBD sequence conservation based on an alignment of 272 reference CP2 motifs in the Grh family from the *Pfam* database. Highest sequence conservation, represented by the widest tube, is observed near the target DNA. (*D*) Sequence conservation between human Grhl1, -2 and -3. Specific sites of protein:DNA-base contacts are marked by asterisks. Residues within the green box are without electron density in the crystal structure and may be involved in additional contacts of Grhl1 with flanking DNA regions not observed in the Grhl1-DBD:DNA crystal structure. TT, tight β turn. Amino-acid numbers are for Grhl1.

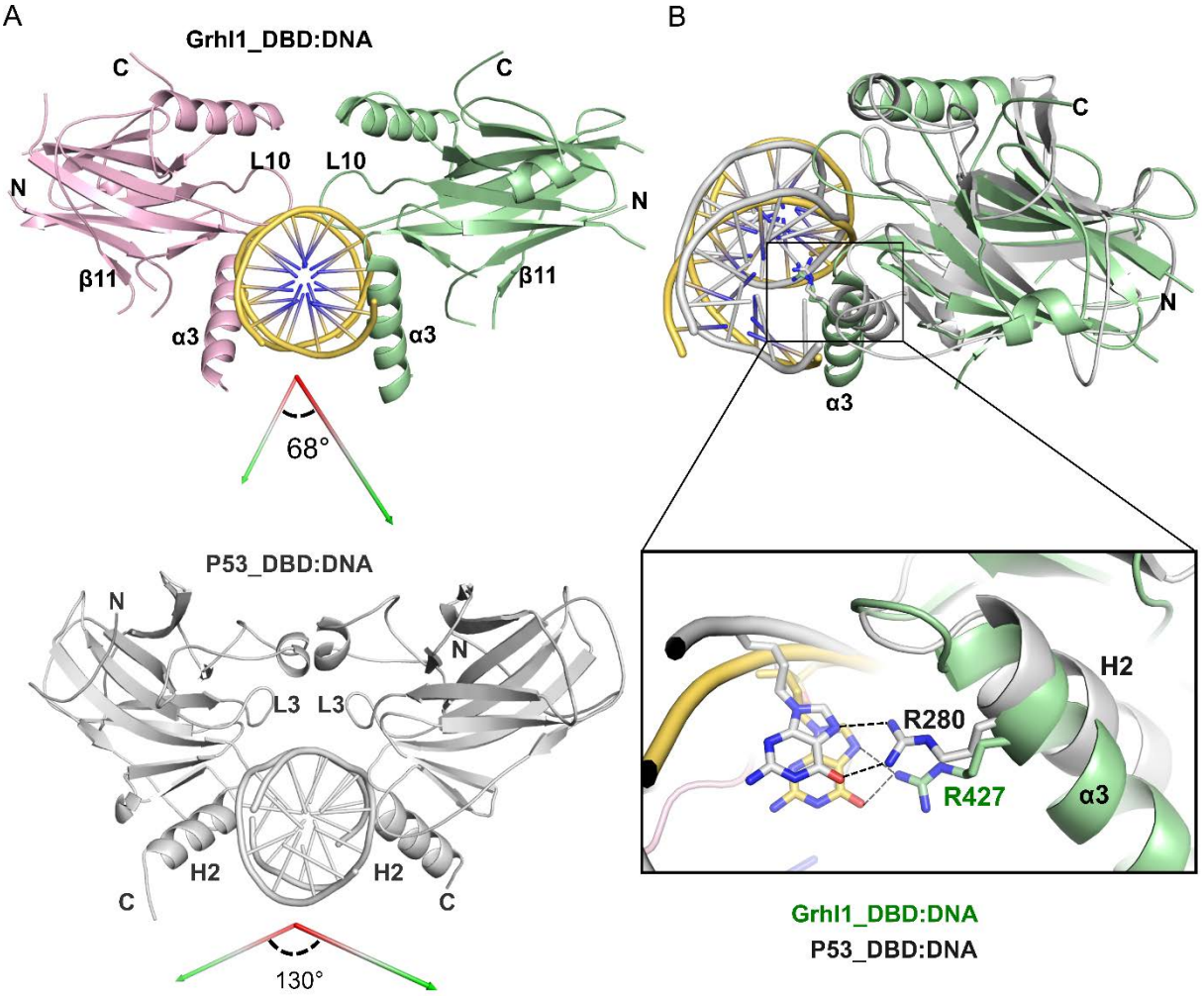



**Fig. S4.** Comparison of DBD:DNA complex structures of Grhl1 and p53 reveals an architecturally similar DNA interacting motif. (*A*) Grhl1-DBD:DNA complex (colored as in Fig. 2*A*) and p53-DBD:DNA complex (gray, PDB code: 2ATA) form similar 2:1 complexes with target DNA. Angles between the two helices bound to the major groove of DNA are different for Grhl1 and p53 as indicated below each complex structure. Inter-helical angles were calculated with Pymol. (*B*) Structural superimposition of DNA-bound Grhl1 DBD (chain B, green) and p53 DBD (chain B, gray) in Pymol. The specific recognition in both structures through an arginine (Arg280 of p53, Arg427 of Grhl1)-guanine interaction is illustrated in the bottom panel.

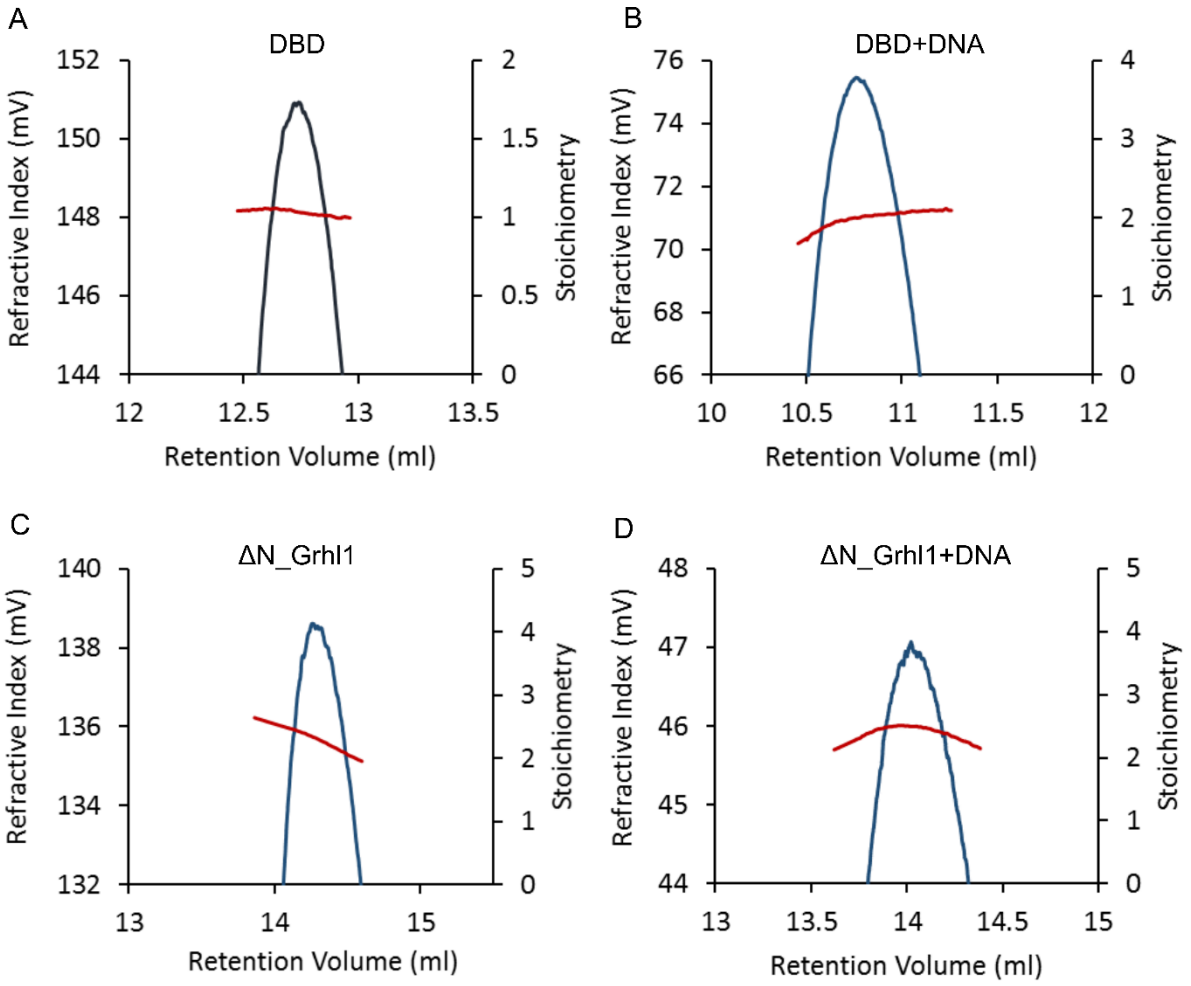



**Fig. S5.** Oligomerization of recombinant Grhl1 proteins as detected by right-angle light scattering (RALS) coupled to size-exclusion chromatography. Grhl1 DBD (*A*) and the Grhl1-DBD:DNA complex (*B*) are analyzed by gel filtration on a Superdex 75 column, ΔN_Grhl1 (*C*) and the ΔN_Grhl1:DNA complex (*D*) are analyzed on a Superdex 200 column. Cons_12mer DNA is used in (*B*) and (*D*). Grhl1 DBD runs as a monomer and dimerizes in presence of DNA. ΔN_Grhl1, containing the C-terminal dimerization domain, is dimeric even without DNA. ΔN_Grhl1 stays dimeric when bound to DNA, suggesting that full-length Grhl1 is dimeric with or without bound DNA.

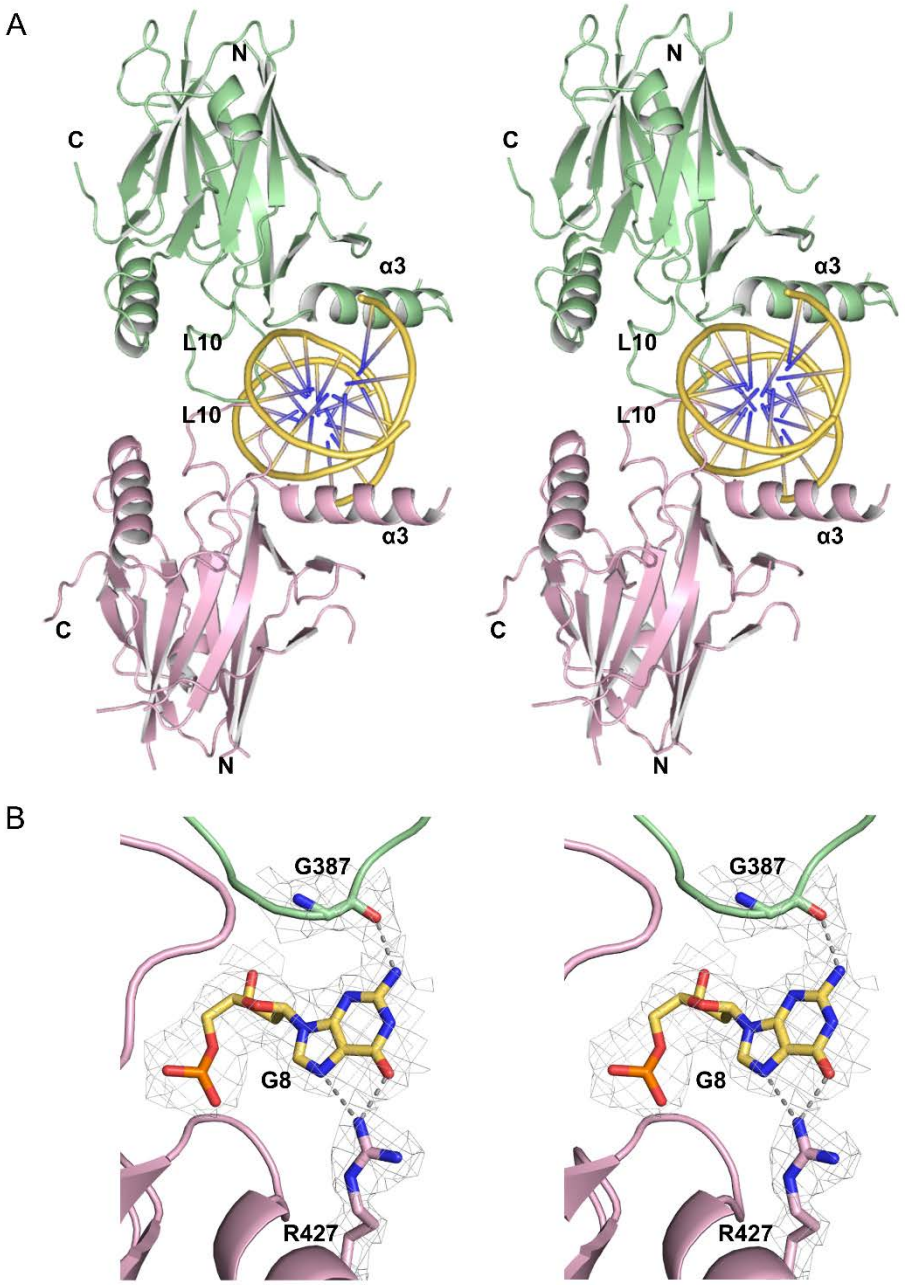



**Fig. S6.** Stereo views of the Grhl1-DBD:DNA complex. (*A*) The overall structure of the DBD:DNA complex. (*B*) Arg427-G8-Gly387 contacts with partial $2F_o$-$F_c$ electron density contoured at 0.8 σ. The protein chains are colored as in Fig. 2.

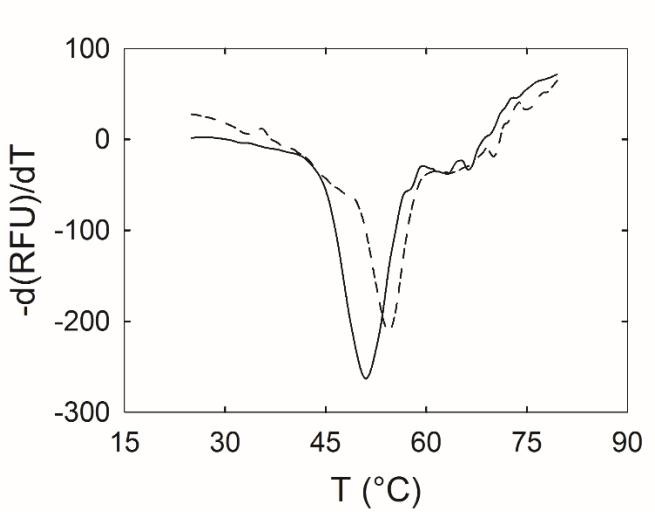



**Fig. S7.** Thermodynamic stability determination. The melting curves for Grhl1 DBD (solid line) and Grhl2 DBD (dashed line) are depicted as the first derivative of the fluorescence intensity against the temperature. The Tm values are 51.0 °C for Grhl1 DBD and 54.5 °C for Grhl2 DBD.

**SUPPLEMENTARY TABLES**

**Table S1.** Mutagenesis primers for reporter gene assay

| Mutation | Forward primer sequence 5'→ 3' | Reverse primer sequence 5'→ 3' |
|---|---|---|
| GRHL1 R427Q | tgacaagggagctgagcagaaaatcagggatgaag | cttcatccctgattttctgctcagctcccttgtca |
| GRHL1 R427A | ctgtgacaagggagctgaggccaaaatcagggatgaagaac | gttcttcatccctgattttggcctcagctcccttgtcacag |
| GRHL2 R423Q | cttctgtgacaaaggagcagaacagaaaatccgagatgaagagcgga | tccgctcttcatctcggattttctgttctgctcctttgtcacagaag |
| GRHL2 R423A | ttctgtgacaaaggagcagaagccaaaatccgagatgaagagcgg | ccgctcttcatctcggattttggcttctgctcctttgtcacagaa |

**Table S2.** ITC measurements of DNA target-site binding by Grhl1 proteins

| Grhl1 | Stoichiometry (N) | $K_d$ (nM) |
|---|---|---|
| DBD | 1.85 ± 0.004 | 90.9 ± 4.5 |
| L378A | 1.77 ± 0.028 | 463 ± 82 |
| T380A | 1.84 ± 0.017 | 521 ± 53 |
| Q385A | 2.03 ± 0.03 | 318 ± 55 |
| K386A | 1.73 ± 0.03 | 1553 ± 155 |
| C421A | 1.61 ± 0.009 | 90.9 ± 1.4 |
| K428A | 1.45 ± 0.003 | 200 ± 6.6 |
| R430A | 2.05 ± 0.02 | 247 ± 33 |
| R427A | 2.03 ± 0.02 | 1,070 ± 67 |
| R427Q | n.d. | > 10,000 |
| R427A/R430A | n.d. | > 10,000 |

**Table S3.** DNA oligonucleotides used for EMSA

| DNA* | Sequence (5'-3') |
|---|---|
| Cy5_12mer | cy5-5'-AAAACCGGTTTT-3' |
| Cons_12mer | 5'-AAAACCGGTTTT-3' |
| Mut_12mer | 5'-AAAATCGATTTT-3' |

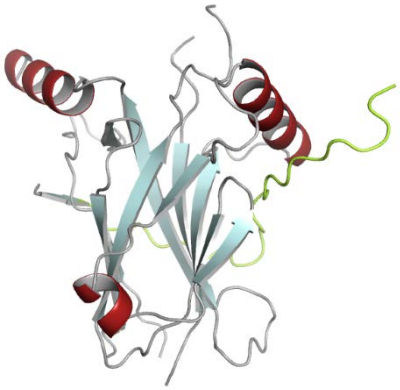
* 12bp DNA duplexes, only one strand shown

**Table S4.** DNA sequence selectivity of Grhl1 DBD

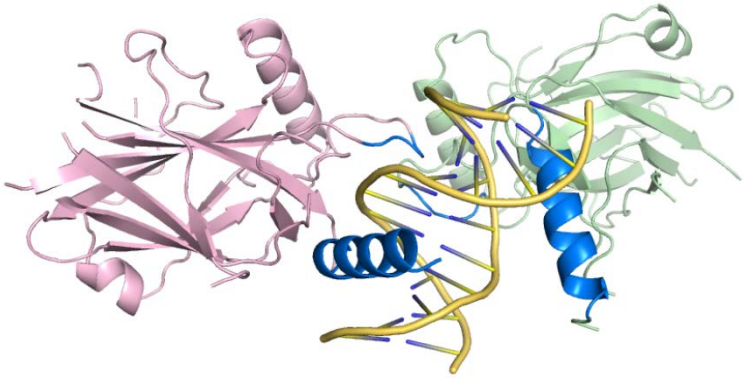
| DNA* | Sequence (5'-3') | Stoichiometry (N) | $K_d$ (nM) |
|---|---|---|---|
| Cons_12mer | AAAACCGGTTTT | 1.85 ± 0.004 | 90.9 ± 4.5 |
| 12AT | AAAACATGTTTT | 1.85 ± 0.02 | 575 ± 42 |
| 12TA | AAAACTAGTTTT | 1.91 ± 0.02 | 493 ± 42 |
| 12AA | AAAACAAGTTTT | 1.64 ± 0.02 | 226 ± 33 |
| 12CC | AAAACCCGTTTT | 1.54 ± 0.01 | 75.8 ± 11 |
| Mut_12mer | AAAATCGATTTT | No binding | No binding |

* 12bp DNA duplexes, only one strand shown

**SUPPLEMENTARY MOVIES**



**Movie S1.** The structure of the Grhl DBD in 360° rotation. The C-terminal peptide has no regular secondary structure and is highlighted in light green.



**Movie S2.** The Grhl1-DBD:DNA binding surface is mainly composed of α3 and L10 interacting with the major and minor groove of DNA, respectively. This region of Grhl1 DBD is positively charged as indicated by the electrostatic surface.