SUPPLEMENTARY DATA

Model-guided mutagenesis drives functional studies of human NHA2, implicated in hypertension

Maya Schushan‡, Minghui Xiang‡, Pavel Bogomiakov, Etana Padan, Rajini Rao* and Nir Ben-Tal*

‡These authors contributed equally to this work.

*Corresponding authors

SUPPLEMENTARY TEXT

Expression and Localization Analysis of NHA2 Mutants

We assessed the membrane localization of the non-functional mutants detected in this study to investigate whether impaired functional phenotypes resulted from miss-localization or expression of the mutant proteins (Supplementary Figs. S4 and S5). Interestingly, all non-functional mutants showed expression and localization to the plasma membrane, suggesting that the impediments indeed resulted from interference of the protein's function and not from defects in its structural organization.

Additional MSAs produced for NHA2

We conducted a three-iteration PSI-BLAST[1] against the UniProt database[2] in search of homologous sequences for the membrane domain of NHA2. Since over 1000 hits were found, we created several collections of homologues, consisting of the 200, 300 and 500 best hits from the third PSI-BLAST round. The sequences chosen for these collection were not fragments (defined as hits shorter than 80% the length of NHA2) or redundant sequences (<90%). After aligning the sequences using the MUSCLE algorithm,[3] we utilized the different MSAs to calculate evolutionary conservation scores via the ConSurf webserver (http://consurf.tau.ac.il,[4]). We then compared the scores received for the different sets to the scores computed for the initial, more limited sequence collection, consisting of only 55 sequences. Reassuringly, mapping of the scores on the model-structure of NHA2 showed generally high correspondence between the different collections, with the protein core remaining conserved in all examined sets, and the periphery of the protein variable. This indicated that our conservation analysis is

robust to the choice of homologous sequences. It is noteworthy that the five charged residues in and near the TM4-TM11 assembly, namely Glu215, Glu278, Glu279, Arg432 and Lys460, receive the highest conservation grade (9) using all the different sequence collections for computing the scores.

SNPs reside in variable regions

According to the dbSNP database,[5] three non-synonymous SNPs have been identified in NHA2. The polymorphisms result in the following amino acids substitutions: I159T, V161A and F357C. The model predicted that all replacements occur in residues situated in extra-cellular regions and display low evolutionary conservation. Ile159 and Val161 are located in the region preceding TM2, while Phe357 resides in the TM7-TM8 loop. Moreover, the predicted extra-cellular location might explain the fact that two of the SNPs (in positions 159 and 357) result in substitutions of hydrophobic amino acids with polar ones. Substitutions of this kind might not be tolerable in the hydrophobic TM domain, especially in lipid-exposed positions (Supplementary Fig. S3b).

SUPPLEMENTARY FIGURE LEGENDS

Fig. S1. Conserved hydrophilic residues in the protein core. In all panels, evolutionarily conserved hydrophilic residues in the TM domain are shown as spheres. Panels A and C depict the protein from within the membrane, with the intra-cellular side below. Panels B and D provide a top-down view of the transporter from the extra-cellular side of the membrane. **(a)-(b)** the 3D model of NHA2 is colored by hydrophobicity according the scale of Kessel and Ben-Tal[6] with blue-through-yellow corresponding to hydrophilic-through-hydrophobic. **(c)-(d)** The model-structure is colored by conservation scores, using the color-coding bar, with turquoise-through-maroon indicating variable-through-conserved. It is apparent that most of the conserved hydrophilic residues of the TM domain reside in (or in the vicinity of) the TM4-TM11 assembly, or within the external and internal funnels.

Fig. S2. The pairwise alignment between NHA2 and NhaA utilized for building the NHA2 model-structure. The pairwise alignment of the TM domains of NhaA and NHA2, produced using the NHA2 TM boundaries of Fig. 1. NHA2 amino acids that are identical to their corresponding residues in NhaA are highlighted in black and similar residues in grey background. The TM helices are marked by brown rectangles and roman numbers.

Fig. S3. Model validation via the 'positive-inside' rule and SNPs. The intra-cellular side is at the bottom**.** The approximate boundaries of the hydrocarbon region of the membrane, of 25Å width, are shown in grey. **(a)** The NHA2 3D-model is colored white. Cα atoms of lysines and arginines positioned in the intra- and extra-cellular ends of NHA2 are shown as blue spheres. In correspondence with the 'positive-inside' rule,[7] the intra-cellular end contains 15 positively charged residues, while the extra-cellular end includes only eight. **(b)** The model-structure is colored by evolutionary conservation as in

Fig. 2 and according to the color bar below. The three known SNPs, I59, V161 and F359 are shown as spheres. The examined polymorphisms map to variable residues, located on loop regions.

Fig. S4. Expression level of NHA2 mutants heterologously expressed in yeast strain AB11c. Western blot of total lysate (100 µg) from yeast harboring NHA2 (WT), empty vector (EV) or NHA2 mutant, as indicated. Membranes were probed with anti-NHA2 and anti-actin.

Fig. S5. NHA2 mutants localize correctly to the plasma membrane. Fluorescence images were taken from overnight grown AB11c yeasts expressing GFP-NHA2 (WT) or its mutant as indicated.

Fig. S6. Functional analysis of additional NHA2 mutants: dosage response. The salt-sensitive yeast strain AB11C was transformed with His-tagged NHA2 (WT), NHA2 mutant as indicated, or empty vector (EV). Yeasts were grown in APG medium[8] supplemented with LiCl (Left panels), or NaCl (Right panels), and growth was determined by optical density of the culture at 600 nm (OD600) after 24 hrs (LiCl) or 48 hrs (NaCl) at 30° C. Data are the average of triplicate determinations. All mutants were with His tag except mutants in panel B, where NHA2 and its mutants E381C, K382C, E384C, E386C and K387C were tagged with a version of GFP with potential use of pH measurements.

Fig. S7. Functional analysis of additional NHA2 mutants: pH profile. Same as Supplemental Fig. S6 except that LiCl concentration was fixed at 25 mM and NaCl concentration at 300 mM. Media pH was adjusted by using phosphoric acid.

SUPPLEMENTARY FIGURES

Fig. S1

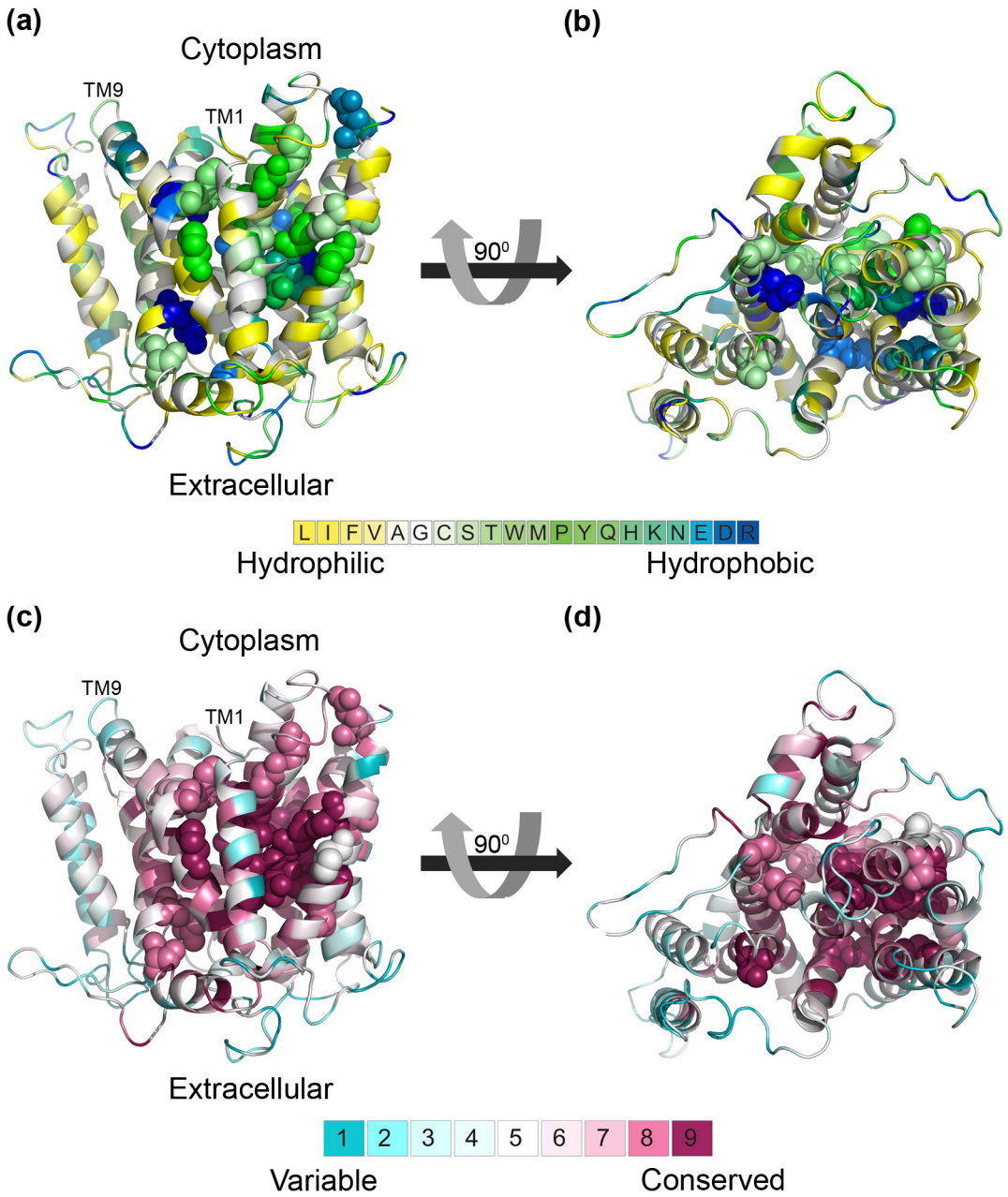


Fig. S2

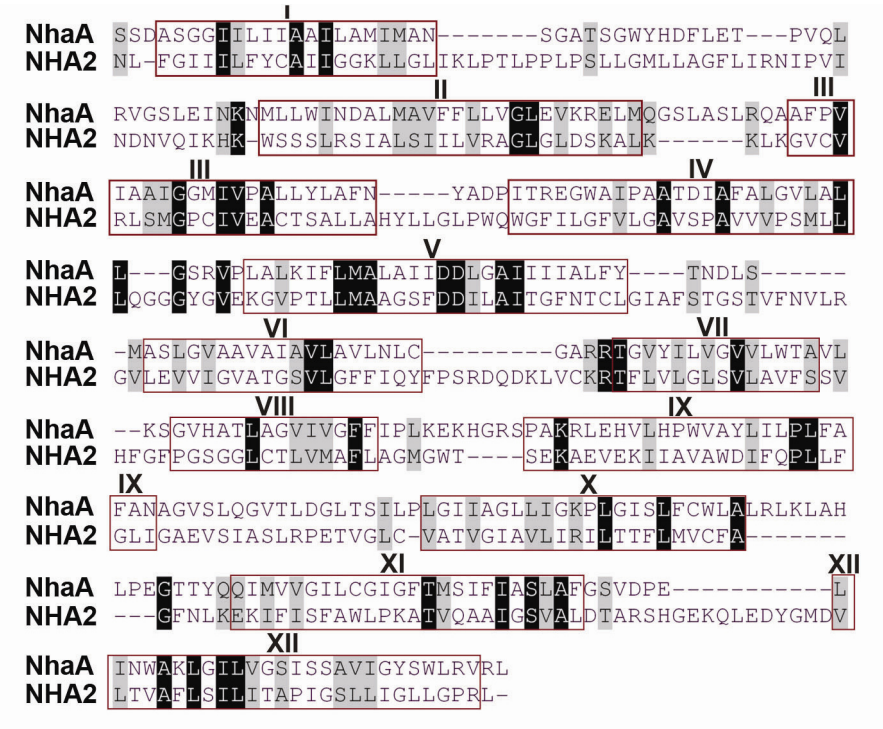

I
NhaA SSDASGGIILIIAAILAMIMAN-------SGATSGWYHDFLET---PVQL
NHA2 NL-FGIIILFYCAIIGGKLLGLIKLPTLPPLPSLLGMLLAGFLIRNIPVI
II
III
NhaA RVGSLEINKNMLLWINDALMAVFFLLVGLEVKRELMQGSLASLRQAAFPV
NHA2 NDNVQIKHK-WSSSLRSIALSIILVRAGLGLDSKALK------KLKGVCV
III
IV
NhaA IAAIGGMIVPALLYLAFN-----YADPITREGWAIPAATDIAFALGVLAL
NHA2 RLSMGPCIVEACTSALLAHYLLGLPWQWGFILGFVLGAVSPAVVVPSMLL
V
NhaA L---GSRVPLALKIFLMALAIIDDLGAIIIIALFY----TNDLS------
NHA2 LQGGGYGVEKGVPTLLMAAGSFDDILAITGFNTCLGIAFSTGSTVFNVLR
VI
VII
NhaA -MASLGVAAVAIAVLAVLNLC---------GARRTGVYILVGVVLWTAVL
NHA2 GVLEVVIGVATGSVLGFFIQYFPSRDQDKLVCKRTFLVLGLSVLAVFSSV
VIII
IX
NhaA --KSGVHATLAGVIVGFFIPLKEKHGRSPAKRLEHVLHPWVAYLILPLFA
NHA2 HFGFPGSGGLCTLVMAFLAGMGWT----SEKAEVEKIIAVAWDIFQPLLF
IX
X
NhaA FANAGVSLQGVTLDGLTSILPLGIIAGLLIGKPLGISLFCWLALRLKLAH
NHA2 GLIGAEVSIASLRPETVGLC-VATVGIAVLIRILTTFLMVCFA-------
XI
XII
NhaA LPEGTTYQQIMVVGILCGIGFTMSIFIASLAFGSVDPE----------L
NHA2 ---GFNLKEKIFISFAWLPKATVQAAIGSVALDTARSHGEKQLEDYGMDV
XII
NhaA INWAKLGILVGSISSAVIGYSWLRVRL
NHA2 LTVAFLSILITAPIGSLLIGLLGPRL-

Fig. S3

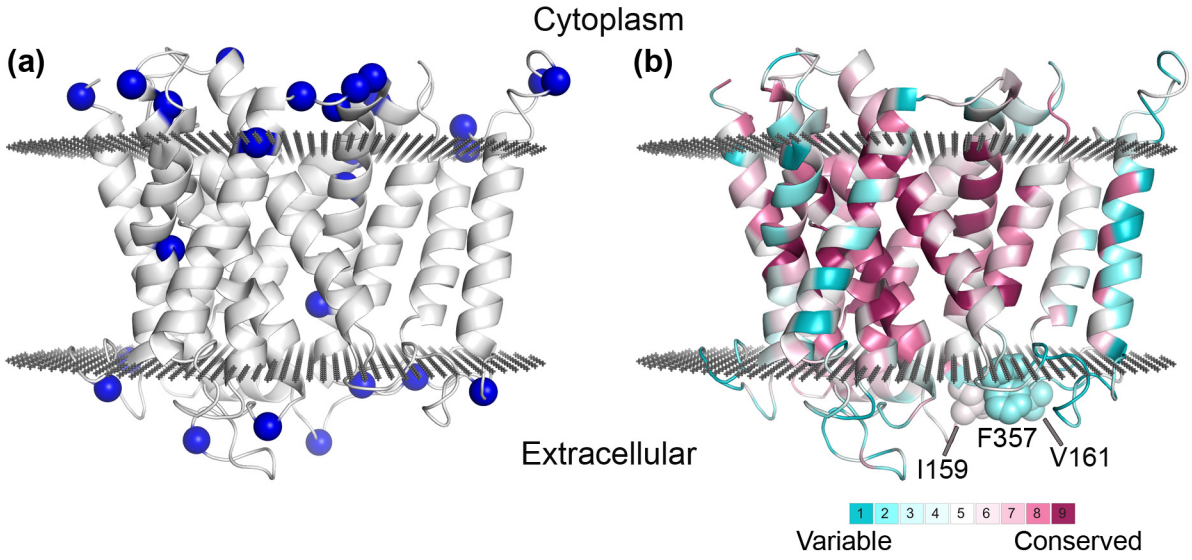

Cytoplasm
(a)
(b)
Extracellular
F357
V161
I159
1 2 3 4 5 6 7 8 9
Variable
Conserved

Fig. S4

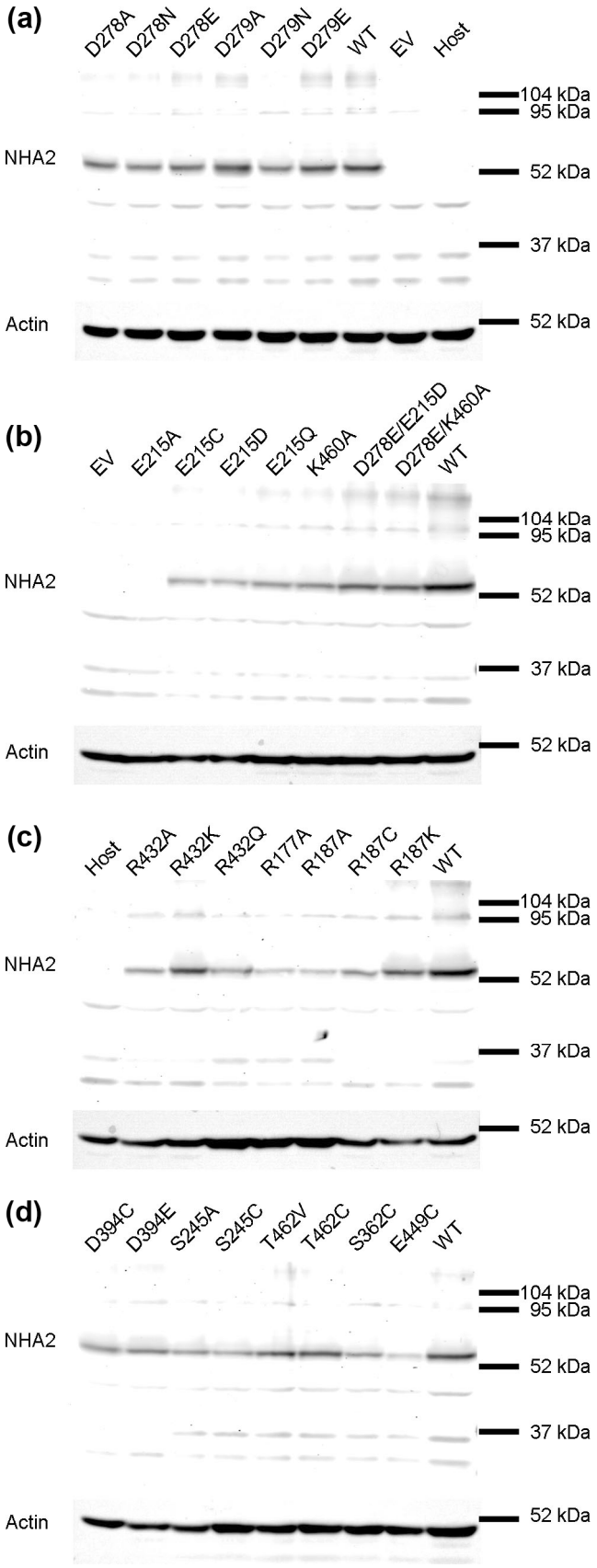

(a)
D278A
D278N
D278E
D279A
D279N
D279E
WT
EV
Host
104 kDa
95 kDa
NHA2
52 kDa
37 kDa
Actin
52 kDa
(b)
EV
E215A
E215C
E215D
E215Q
K460A
D278E/E215D
D278E/K460A
WT
104 kDa
95 kDa
NHA2
52 kDa
37 kDa
Actin
52 kDa
(c)
Host
R432A
R432K
R432Q
R177A
R187A
R187C
R187K
WT
104 kDa
95 kDa
NHA2
52 kDa
37 kDa
Actin
52 kDa
(d)
D394C
D394E
S245A
S245C
T462V
T462C
S362C
E449C
WT
104 kDa
95 kDa
NHA2
52 kDa
37 kDa
Actin
52 kDa

Fig. S5

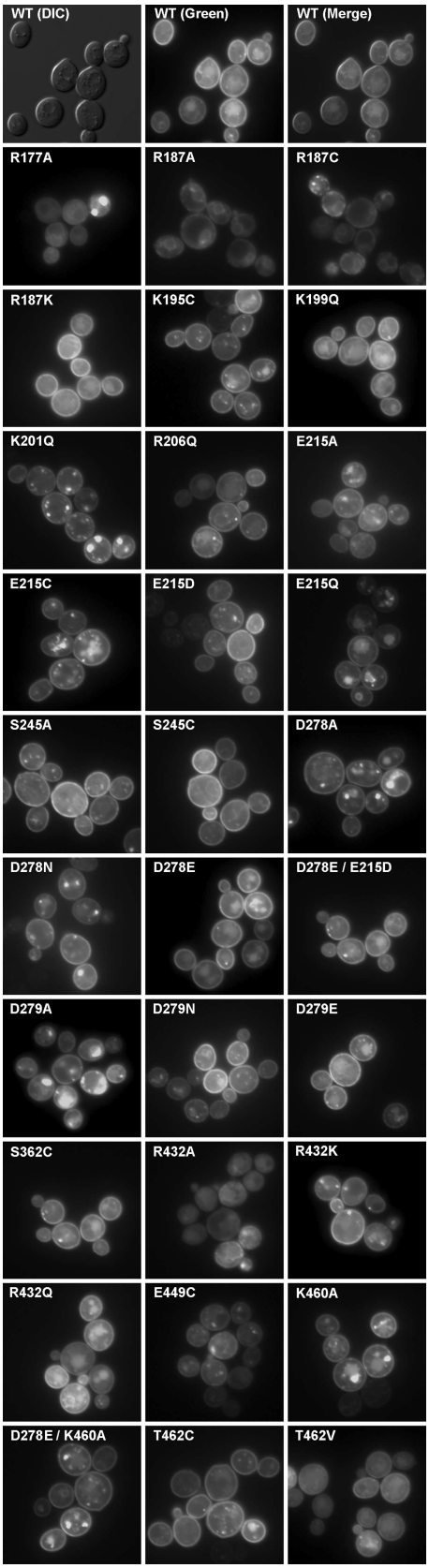

WT (DIC)
WT (Green)
WT (Merge)
R177A
R187A
R187C
R187K
K195C
K199Q
K201Q
R206Q
E215A
E215C
E215D
E215Q
S245A
S245C
D278A
D278N
D278E
D278E / E215D
D279A
D279N
D279E
S362C
R432A
R432K
R432Q
E449C
K460A
D278E / K460A
T462C
T462V

Fig. S6

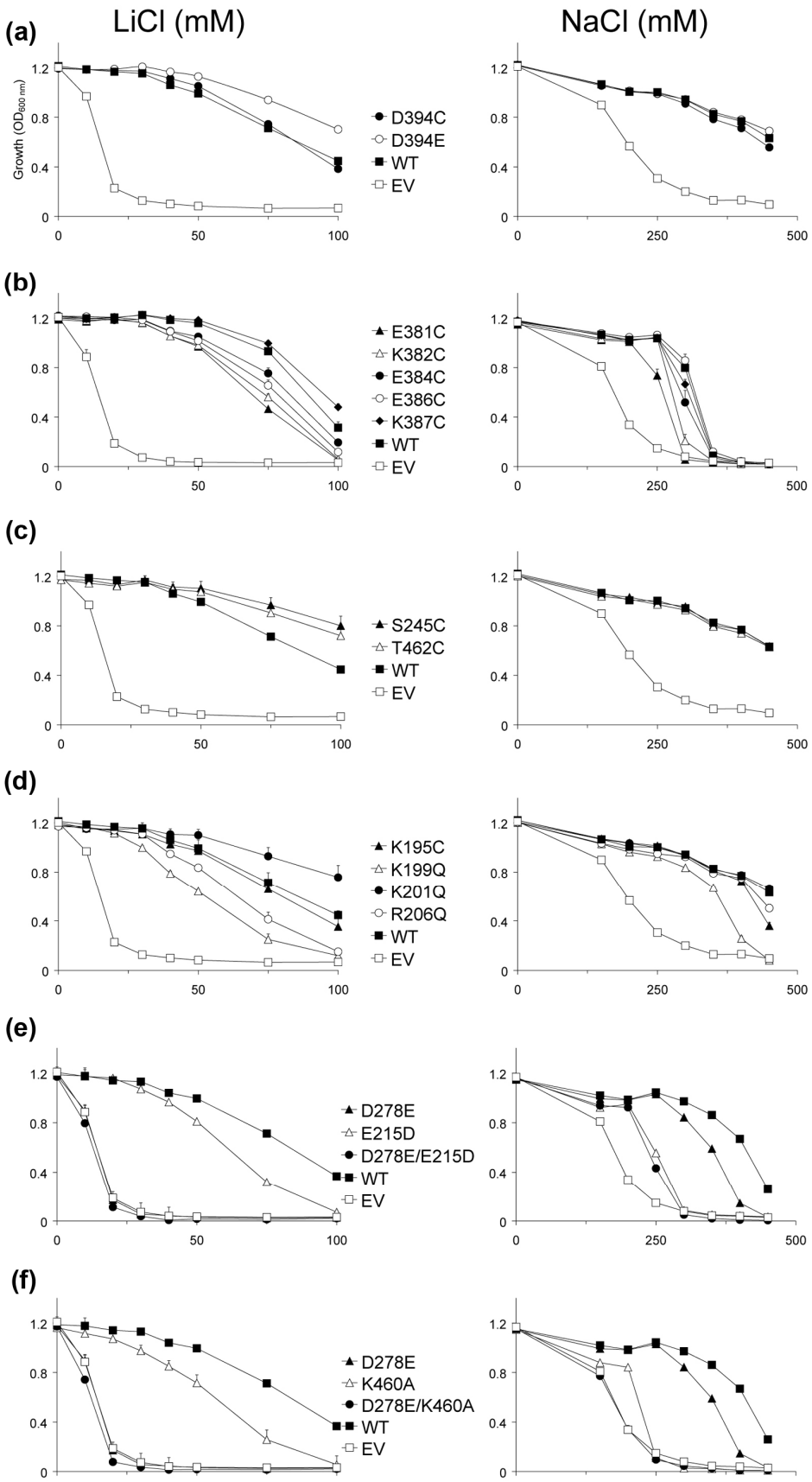
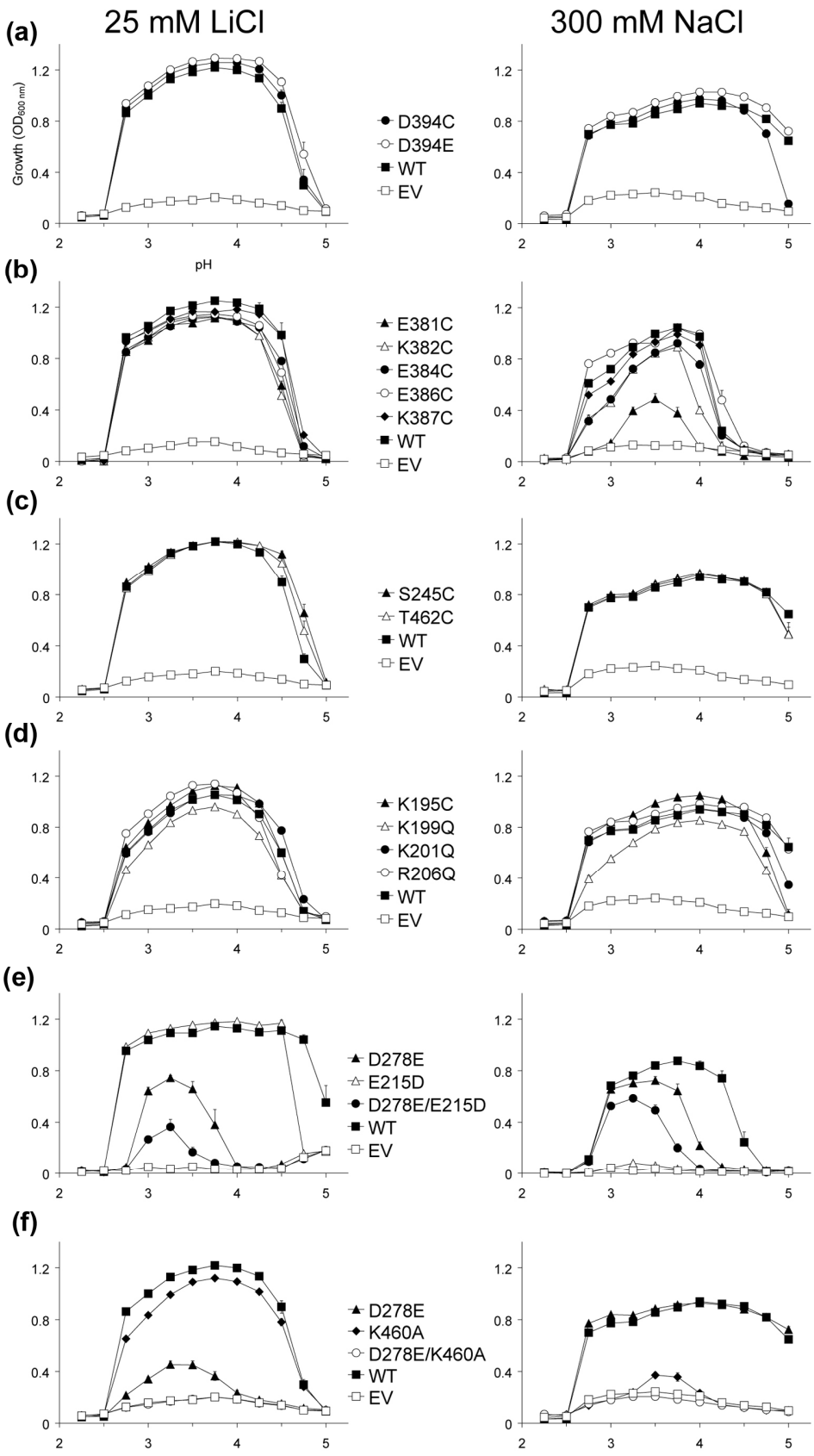

LiCl (mM)
NaCl (mM)
(a)
Growth ($OD_{600\,nm}$)
D394C
D394E
WT
EV
(b)
E381C
K382C
E384C
E386C
K387C
WT
EV
(c)
S245C
T462C
WT
EV
(d)
K195C
K199Q
K201Q
R206Q
WT
EV
(e)
D278E
E215D
D278E/E215D
WT
EV
(f)
D278E
K460A
D278E/K460A
WT
EV

Fig. S7


25 mM LiCl
300 mM NaCl
(a)
(b)
(c)
(d)
(e)
(f)
Growth ($OD_{600\,nm}$)
pH
D394C
D394E
WT
EV
E381C
K382C
E384C
E386C
K387C
S245C
T462C
K195C
K199Q
K201Q
R206Q
D278E
E215D
D278E/E215D
K460A
D278E/K460A

References

1. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, pp. 3389-3402.
2. Bairoch, A., Apweiler, R., Wu, C. H., Barker, W. C., Boeckmann, B., Ferro, S., Gasteiger, E., Huang, H., Lopez, R., Magrane, M., Martin, M. J., Natale, D. A., O'Donovan, C., Redaschi, N. & Yeh, L. S. (2005). The Universal Protein Resource (UniProt). *Nucleic Acids Res* **33**, pp. D154-159.
3. Edgar, R. C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**, pp. 113.
4. Landau, M., Mayrose, I., Rosenberg, Y., Glaser, F., Martz, E., Pupko, T. & Ben-Tal, N. (2005). ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. *Nucleic Acids Res* **33**, pp. W299-302.
5. Sherry, S. T., Ward, M. H., Kholodov, M., Baker, J., Phan, L., Smigielski, E. M. & Sirotkin, K. (2001). dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* **29**, pp. 308-311.
6. Kessel, A. & Ben-Tal, N. (2002). Free energy determinants of peptide association with lipid bilayers. In *Current Topics in Membranes 52* (Simon S.A. & McIntosh T. J.), pp. 205-253, Academic Press, San Diago.
7. Heijne, G. V. (1986). The distribution of positively charged residues in bacterial inner membrane proteins correlates with the trans-membrane topology. *EMBO J* **5**, pp. 3021-3027.
8. Rodriguez-Navarro, A. & Ramos, J. (1984). Dual system for potassium transport in Saccharomyces cerevisiae. *J Bacteriol.* **159**, pp. 940-945.