

patchWork: a user-friendly pH sensitivity analysis web server for protein sequences and structures

Mirko Schmitz^{1,2}, Anne Schultze^{1,2}, Raimonds Vanags³, Karsten Voigt³,
Barbara Di Ventura^{1,2} and Mehmet Ali Öztürk^{1,2,*}

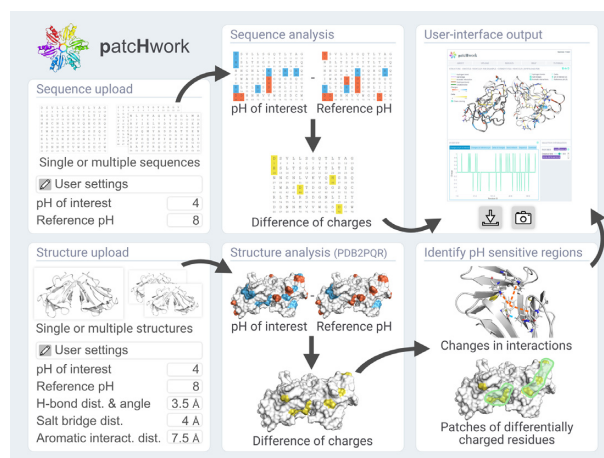
¹Signaling Research Centres BIOS and CIBSS, University of Freiburg, Schänzlestr. 18, 79104 Freiburg, Germany, ²Institute of Biology II, University of Freiburg, Schänzlestr. 1, 79104 Freiburg, Germany and ³Institute of Biology III, University of Freiburg, Schänzlestr. 1, 79104 Freiburg, Germany

Received February 06, 2022; Revised March 11, 2022; Editorial Decision March 29, 2022; Accepted April 12, 2022

ABSTRACT

pH regulates protein function and interactions by altering the charge of individual residues causing loss or gain of intramolecular noncovalent bonds, which may lead to structural rearrangements. While tools to analyze residue-specific charge distribution of proteins at a given pH exist, currently no tool is available to investigate noncovalent bond changes at two different pH values. To make protein pH sensitivity analysis more accessible, we developed patchWork, a web server that combines the identification of amino acids undergoing a charge shift with the determination of affected noncovalent bonds at two user-defined pH values. At the sequence-only level, patchWork applies the Henderson–Hasselbalch equation to determine pH-sensitive residues. When the 3D protein structure is available, patchWork can be employed to gain mechanistic understanding of the effect of pH. This is achieved using the PDB2PQR and PROPKA tools and noncovalent bond determination algorithms. A user-friendly interface allows visualizing pH-sensitive residues, affected salt bridges, hydrogen bonds and aromatic (π – π and cation– π) interactions. patchWork can be used to identify patches, a new concept we propose of pH-sensitive residues in close proximity on the protein, which may have a major impact on function. We demonstrate the attractiveness of patchWork studying experimentally investigated pH-sensitive proteins (<https://patchwork.biologie.uni-freiburg.de/>).

GRAPHICAL ABSTRACT



INTRODUCTION

The concentration of hydrogen ions in a solution (used to calculate the pH) determines the charge of the side chains of the amino acids in a given protein by regulating their protonation state. Various properties such as protein solubility (1), stability (2), ability to interact with other molecules (3), flexibility (4) and activity (5) are affected by pH. As a matter of fact, the amount of protonation of the amino acid side chains in proteins has been proposed to be a new form of protein post-translational modification (6). Ionizable residues in proteins react to the surrounding pH according to their ionization constant (pK_a) and they become positively or negatively charged at pH values below or above their pK_a , respectively. This adaptation can cause loss or gain of intramolecular noncovalent bonds between residues, which can subsequently result in structural rearrangements that could regulate protein function and interactions (7,8). Thus, it is crucial to understand the effect of pH change on a protein structure and determine any resulting adjustments in intramolecular noncovalent bonds.

*To whom correspondence should be addressed. Tel: +49 761 2032787; Email: mehmet.oeztuerk@bioss.uni-freiburg.de

Currently, a number of web servers exist that allow calculating the charges of the amino acid side chains at a certain pH using either sequence or structural information (9–11). VOLPES (10), for instance, applies the Henderson–Hasselbalch equation (12) to calculate the charges of the side chains of the amino acids in a protein of interest at a user-defined pH using only sequence information. However, the Henderson–Hasselbalch equation does not account for the influence of neighboring amino acids on the pK_a of a residue side chain (13). An analysis at the level of the structure is, therefore, clearly required to precisely determine the effect of pH change on proteins.

Structure-based analysis of pH-mediated changes in the charge of the side chains of amino acids has benefited from datasets obtained from NMR experiments, which significantly helped refine computational methods for pK_a prediction of residues in protein structures (14). For example, APBS (11) calculates the electrostatic potentials of proteins by assigning charge and radius to atoms using the PDB2PQR (15) and PROPKA (16) tools. Protein-sol (9) provides overall pH-dependent charge information for proteins of interest, as well as predictions of the destabilization of residue-specific electrostatic interactions due to limited ionization ability of buried amino acids (13). There is also a graphical user interface (GUI) plug-in implementation (17) on the VMD software (18) to visualize PROPKA (16) predictions. While being useful, these approaches do not allow the user to easily monitor the appearance/disappearance of intramolecular noncovalent bonds (salt bridges, hydrogen bonds, π – π and cation– π interactions) when the pH is changed from one value to another. On the other hand, these bonds can be investigated only at the default physiological pH with web servers such as RING 2.0 (19), Arpeggio (20) and ProteinTools (21). Taken together, currently no tool offers the possibility to directly observe protonation changes of amino acids caused by a shift in pH between two user-defined values and the resulting gain/loss of noncovalent bonds in the protein structure. Therefore, researchers wishing to analyze pH sensitivity of a given protein need to use several tools in parallel with manual curation of the outputs, which makes such an analysis difficult when the knowledge in computational structural biology is limited.

Here, we present patchHwork, a novel web server that supports high-throughput pH sensitivity analysis at two user-defined pH values at either the sequence or structure level. At the sequence level, patchHwork allows submitting up to 10 000 protein sequences, which are then analyzed using the Henderson–Hasselbalch equation (12) to determine pH-sensitive residues at the user-defined pH values. When the 3D protein structure is available, further mechanistic understanding of the effect of pH on a protein of interest can be obtained by the execution of the PDB2PQR (15) and PROPKA (16) software and noncovalent bond determination algorithms (22–25). pH-sensitive residues and pH-mediated changes in salt bridges, hydrogen bonds and aromatic (π – π and cation– π) interactions are visualized in an interactive GUI. Additionally, users obtain information regarding so-called patches, groups of pH-sensitive residues found in a customizable physical distance on the protein structure, which may play a more profound role than individual amino acids in the regulation of protein function

upon pH change. To demonstrate the workflow and the power of patchHwork, we carried out sequence-based pH sensitivity analysis of *Escherichia coli* cell envelope proteins, and structure-based analysis of the taste-modifying protein neoculin (26,27) and the pH-regulated mouse anion exchanger 2 (mAE2) protein (28–32).

FUNCTIONALITIES OF patchHwork

patchHwork has four main computational components to investigate the response of proteins to a change in pH: protein sequence-based analysis, protein structure-based analysis, noncovalent bond analysis and identification of pH-sensitive patches.

Protein sequence-based analysis

The Henderson–Hasselbalch equation (12) is solved for each amino acid of the submitted protein FASTA sequences at two user-given pH values. Residue-specific charges at each pH value and also delta charges are provided as an interactive output. In order to rank the proteins based on their pH sensitivity, an ‘overall charge score’ is defined as follows: for each protein, the sum of the charges at the pH of interest is subtracted by the sum of charges at the reference pH and then normalized by the total number of residues in the protein.

Protein structure-based and noncovalent bond analyses

The protonation state of each amino acid in the submitted protein PDB structures is calculated using the PDB2PQR (15) and PROPKA (16) tools at the two user-given pH values. In addition to residue-specific charge information, created and destroyed noncovalent bonds (salt bridges, hydrogen bonds, and π – π and cation– π interactions) upon pH change are determined and given as an interactive output.

pH-sensitive patches

On protein structures, residues that change their protonation state at a given pH shift and are found within a radius of ≤ 8 Å from each other (customizable) are defined as a pH-sensitive patch.

Further details of sequence and structure-based analyses, pH-sensitive patch identification and noncovalent bond determination, as well as the web server implementation, are given in Supplementary Data.

CASE STUDIES

Sequence-based pH sensitivity analysis of *E. coli* cell envelope proteins

To demonstrate the advantage of using patchHwork for pH sensitivity analysis, we asked whether we could identify in a high-throughput manner proteins that are mostly affected by pH looking at a family of proteins that is exposed to the extracellular medium and, consequently, is most likely affected by pH change than cytoplasmic proteins: cell envelope proteins. We collected the FASTA sequences of 309 *E. coli* proteins annotated as being part of the cell envelope

(GO:0030313) and determined for each the overall charge score (residues' total charge shifts normalized by the protein length) when changing the pH from 1 to 14 with an increment of 1 (i.e. pH of interest 2, reference pH 1; then, pH of interest 3, reference pH 2; and so on; for details see Supplementary Data). Next, we created for each protein a mean score (\bar{x}) taking the mean of the overall charge scores of all pH shifts (Figure 1A). We propose that, by ranking the proteins based on the \bar{x} score, we are able to identify the most and least pH-responsive proteins, respectively. After obtaining this ranked list, we checked the literature for experimental information on the top and bottom five proteins to verify whether they are reported to be involved in pH regulation/responsiveness or not. Interestingly, while four out of the five top ranked proteins are reported to be pH-responsive/regulated, for the bottom five proteins only one is reported to be regulated by/sensitive to pH (see Supplementary Table S1). When we ran the top-ranked ZinT protein sequence in patchwork at pH values of 6 and 9, we captured the amino acids involved in metal ion binding (residues 24–29 and 166–178), which have been previously implicated in pH sensing (Figure 1B) (33). This example illustrates how, using patchwork, it is possible to conduct quantitative pH sensitivity analysis, ranking hundreds of protein sequences in a matter of minutes.

Structural analysis of pH-induced changes in the sweet taste protein neoculin

To demonstrate how patchwork can be used to narrow down potential residues likely to be involved in the pH response mechanism of a given protein, we analyzed neoculin, a heterodimeric protein from the plant *Curculigo latifolia*. Neoculin (also called curculin) consists of an acidic (NAS) and a basic (NBS) subunit and it turns sour into sweet taste (34). This effect has been shown to be induced by low pH (35). A mutant neoculin where all five histidines were mutated to alanine was shown to be active across pH conditions (26). Specifically, His11 in the NBS was identified as the main pH sensor responsible for the activity of the protein at low pH. Nakajima *et al.* suggested that low pH-mediated loss of the aromatic interaction between His11 and His14 in the NBS is essential for the pH-responsive function of neoculin (26).

While it is intuitive to mutate histidine residues to investigate the pH response mechanism of a given protein considering that the pK_a of histidine is close to the physiological pH, analyzing the protein structure could give further valuable information to narrow candidate residues down as well as give insights into a potential mechanism. For this purpose, we ran a structure-based analysis of neoculin (PDB ID: 2D04, chains A and B) (36) with patchwork for the pH shift from 4 to 8, the two extreme pH values used in the experimental study (26) (session can be accessed at <https://patchwork.biologie.uni-freiburg.de/results.php?key=example.pdb>). We found that the noncovalent bonds between Arg38 and His36 (aromatic interaction), His67 and Ser50 (hydrogen bond), Arg53 and His11, and His11 and His14 (aromatic interactions) are destroyed. Interestingly, His11 has two disrupted noncovalent bonds, while the other differentially charged residues only have sin-

gle ones, hinting that a change in pH especially affects this region occurring in the Arg53–His11–His14 triad, which is in line with the experimental observations (26) (Figure 2A and B). In addition to the analysis of noncovalent bond changes, we identified a pH-sensitive patch constituted by His11, His14 and His67 (Figure 2C).

Structural analysis of pH-regulated mAE2

To showcase the usefulness of patchwork for the pH sensitivity analysis of proteins consisting of over a thousand amino acids, for which the shift in the pK_a of residues is due to the surrounding environment, we selected as the last case study an anion exchanger protein (also called bicarbonate transporter).

The anion exchangers 1–3 (AE1–3) mediate Na^+ -independent $\text{Cl}^-/\text{HCO}_3^-$ exchange in several types of cells. Their functions range from gas transport to cell volume and intracellular pH regulation (37). Both AE1 and AE2 catalyze $\text{H}^+ - \text{SO}_4^{2-} / \text{Cl}^-$ and $\text{H}^+ - \text{SO}_4^{2-} / \text{H}^+ - \text{SO}_4^{2-}$ exchange in a pH-sensitive fashion: in an acidic environment (pH of 5.5), the $\text{H}^+ - \text{SO}_4^{2-}$ efflux is maximal, whereas it sharply decreases at near-neutral pH (pH of 7.5) (28–31). In human AE1 (hAE1), the binding site for the proton that is cotransported with SO_4^{2-} was proposed to be Glu681 (31). Chemically converting this negatively charged residue to an alcohol (Glu681OH), and thus rendering it neutrally charged, leads to proton-independent transport of SO_4^{2-} as well as nearly complete elimination of the transport's pH dependence (32,38). Mutating the corresponding conserved glutamic acid in mouse AE1 (mAE1) and mAE2, Glu699 and Glu1007, respectively, to the neutral glutamine abolishes the preference of sulfate exchange in acidic medium similarly to the experiments with hAE1 (28–32,38).

In the absence of this experimental evidence, we would typically perform mutagenesis analysis on mAE2 to identify amino acids involved in the pH-regulated SO_4^{2-} transport. To this aim, we would likely focus on the histidines, as their pK_a value is close to the physiological pH and thus we know that they would change protonation state for a pH shift from 5.5 to 7.5. This methodology would lead to identifying 35 candidate histidine residues in mAE2, which consists of 1237 amino acids. However, it would not lead to identifying Glu1007, which is reported to be the key pH-responsive amino acid of mAE2 (28–30).

Using patchwork for the analysis allows avoiding falling in the prototypical histidine-oriented approach. As a matter of fact, by making use of pK_a predictions from PROPKA (16), patchwork calculates the shift in pK_a considering the surrounding environment of the side chains of each amino acid, since this is known to exert a role (13). We downloaded the mAE2 model structure from the AlphaFold Protein Structure Database (39,40) (model ID: AF-P13808-F1), removing regions with low model confidence score (<70) with the exception of the linker between the two domains (anion exchanger and cytoplasmic domains), and we submitted the structure to patchwork using pH of interest of 5.5 and reference pH of 7.5, which were similarly used in the experimental studies (29,30). We obtained the following three levels of information (session can be ac-

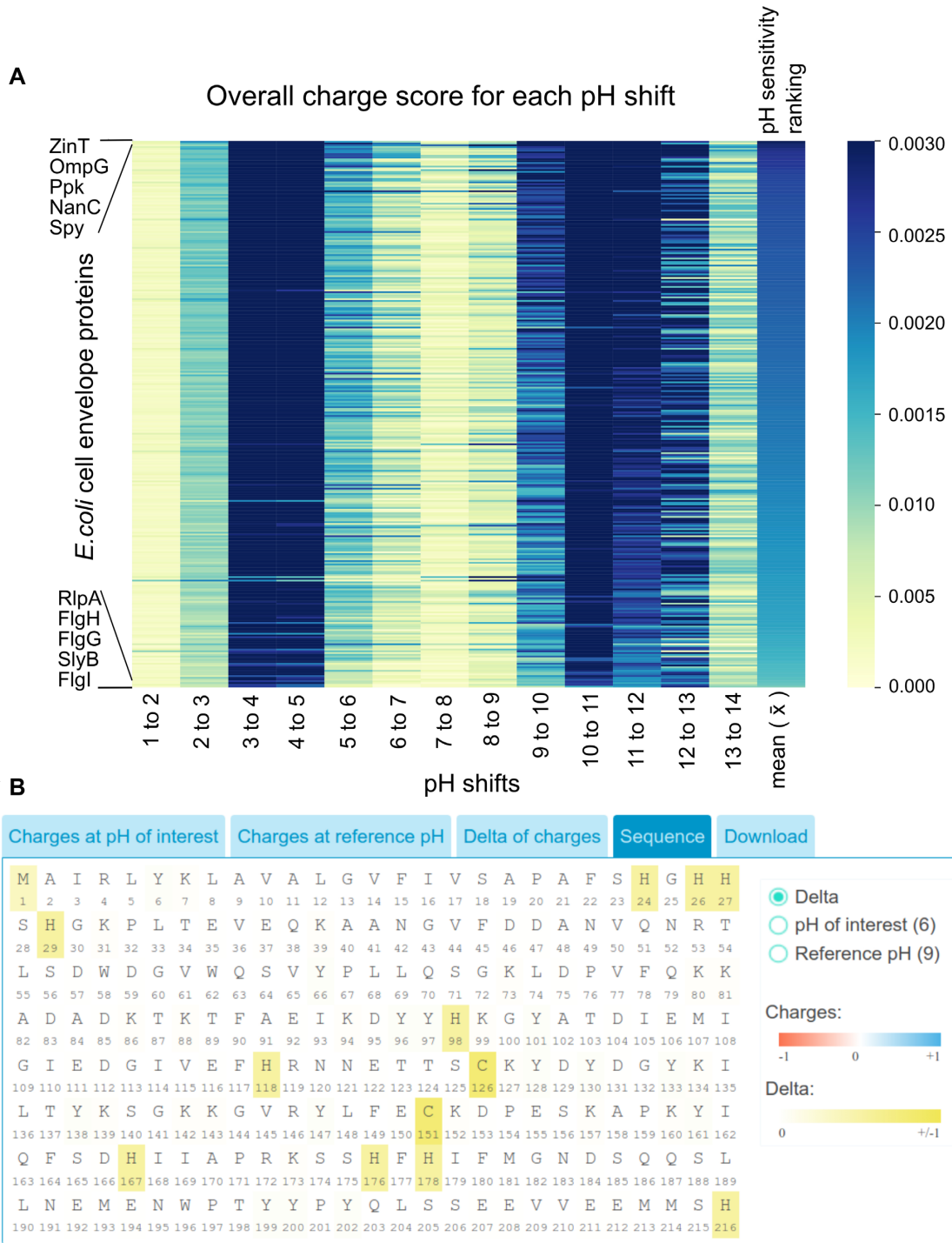


Figure 1. Demonstration of sequence functionality of patchwork. (A) Overall ‘charge score’-based pH sensitivity ranking of 309 *E. coli* cell envelope proteins (GO:0030313) for one unit incremental pH increase from pH 1 to 14. \bar{x} represents the mean of all overall charge scores for each protein. *Note:* Uncharacterized proteins are not considered in Supplementary Table S1. (B) ZinT, the most pH-responsive *E. coli* cell envelope protein from panel (A), is analyzed for the shift in pH from 6 to 9 with patchwork using its sequence functionality.

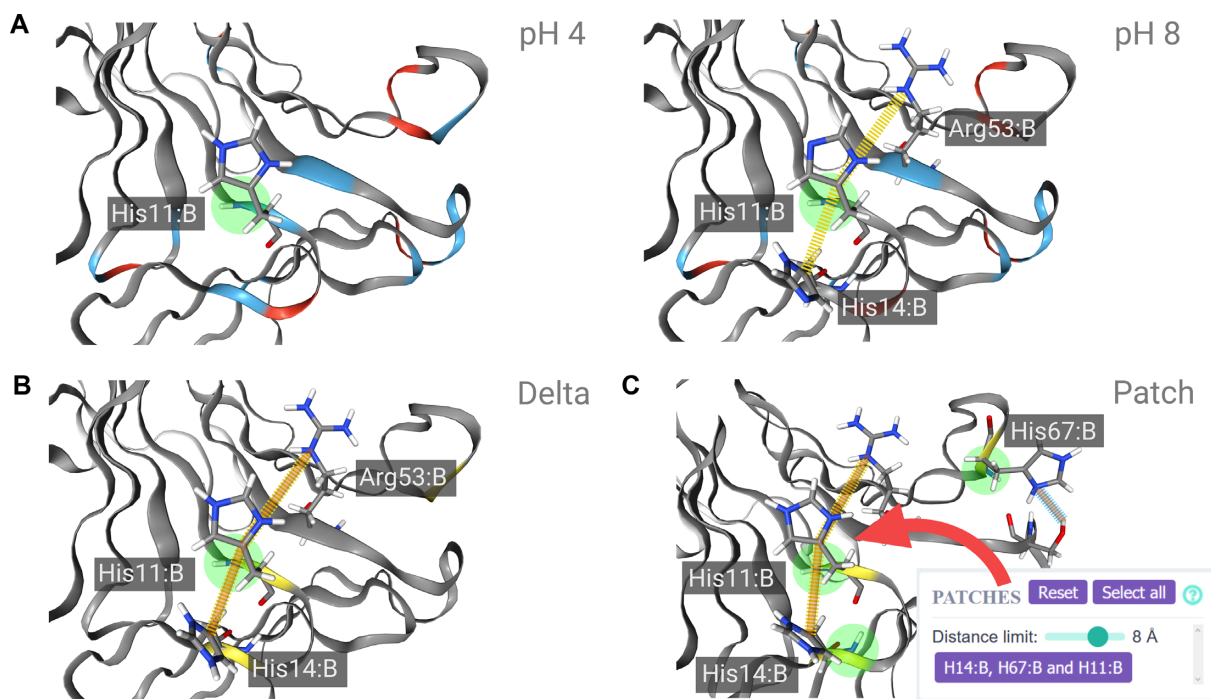


Figure 2. Structure-based patchwork analysis using neoculin X-ray structure (PDB ID: 2D04, chains A and B) at pH values of 4 and 8. (A) The intramolecular pi–pi interaction between His11 and His14 is shown in yellow at pH 8. (B) Delta view in patchwork. Intramolecular noncovalent bonds that are present only for one of the two pH values are shown in orange. (C) Patch view in patchwork. The pH-sensitive patch, identified when the threshold for the distances between pH-sensitive residues is set to be ≤ 8 Å, is shown. Individual residues within the patch are displayed in the box. (A, B) The structure of neoculin with both its subunits (PDB ID: 2D04, chains A and B) was used in the analysis. The figure was generated in patchwork with modifications to the residue labeling. Negatively charged residues are shown in red, positively charged residues in blue and neutral residues in white. Residues that are differentially charged at pH 8 and 4 are highlighted in yellow.

Table 1. Comparison of histidine-based and structure-based analyses (the latter performed with patchwork) for the identification of pH-sensitive amino acids within mAE2

Using mAE2 sequence and pK_a information for pH shift from 5.5 to 7.5	Using mAE2 model structure in patchwork for pH shift from 5.5 to 7.5		
	Protonation state shifting residues	At least one protonation state shifting residue + noncovalent bond change	Part of a patch (≤ 9 Å close to at least two other protonation state shifting residues)
35 histidines	14 histidines Glu338 Glu1007 Asp1031	Asp1031–Gln756 His1029–Asp684 Glu1007–Ser764 Asp684–His1029 His1145–His1160 His360–Arg570 Glu346–His360 His566–Arg570 Glu325–His566 Glu338–Arg341 His395–Arg412	[His425–His566–Glu338–His360] [Asp1031–His1029–Glu1007] [His1145–His1160–His1144–His1136]

cessed at https://patchwork.biologie.uni-freiburg.de/results.php?key=mAE2_example): (i) residues that change their protonation state (shown in yellow in patchwork); (ii) residues that change their protonation state and cause noncovalent bond changes (yellow residues with orange and green stripes in patchwork); and (iii) patches: residues that change their protonation state and are in close physical proximity with other residues that change their protonation state (green spheres in patchwork). These three features generated from patchwork can be used to identify pH-sensitive residues as seen in Table 1.

Using patchwork, it is possible to group pH-sensitive residues into three patches (Figure 3A), resulting in a broader approach compared to using histidine-only information with standard pK_a values. Looking at the structure, we can speculate that a pH-sensitive region influencing substrate binding is more likely to be found within the anion exchanger domain itself, close to the two substrate binding sites [assuming that these are conserved with the sites found in hAE1 (41); see Supplementary Data] rather than within the cytoplasmic domain. In this case, the patch constituted by Asp1031–His1029–Glu1007 is the most likely

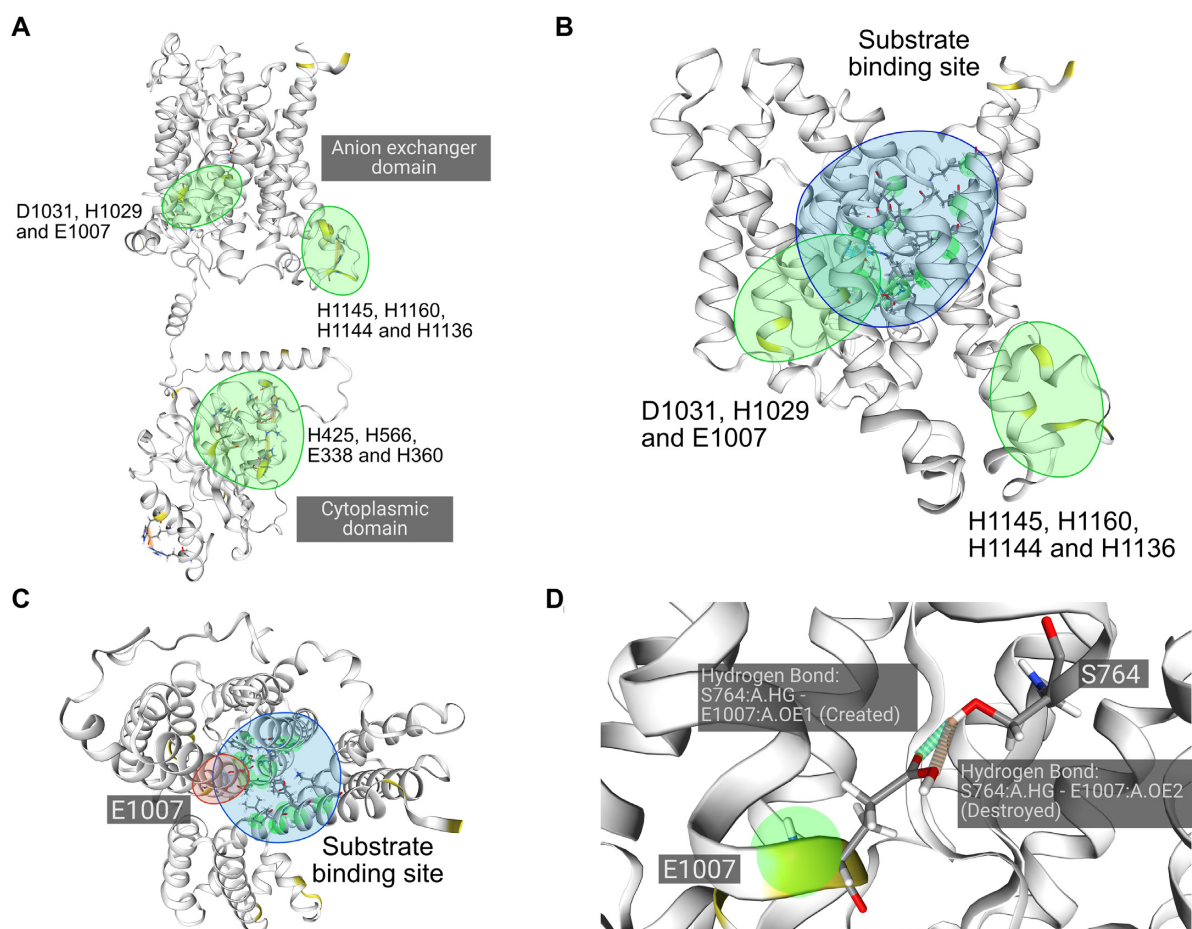


Figure 3. Structure-based patchwork analysis using a model structure of mAE2 at pH values of 5.5 and 7.5. (A) The three pH-sensitive patches identified are highlighted in green. (B) Both patches (highlighted in green) of the anion exchanger domain are shown with the mAE2 substrate binding site (highlighted in blue; for details of the residues, see Supplementary Data). (C) Top view of the substrate binding site and the differentially charged residue Glu1007 closest to it. (D) Shifts in the hydrogen bond between Glu1007 and Ser764, the latter being part of the substrate binding site.

pH-sensitive region of mAE2 (Figure 3B). We conclude that, being the closest to the residues involved in substrate binding, residue Glu1007 is the prime candidate to mutate and experimentally test (Figure 3C). Furthermore, as Glu1007 becomes protonated, a shift in the hydrogen bond with Ser764, belonging to the substrate binding pocket, could hint at a restructuring of the binding pocket allowing for a different substrate such as SO_4^{2-} to bind (Figure 3D). Importantly, our analysis is also in line with an alternative mechanism that was previously proposed, whereby the negatively charged glutamic acid inhibits sulfate binding through ionic repulsion, which ceases when the glutamic acid is protonated (42).

CONCLUSIONS

patchwork is a novel, easy-to-use web server that offers users the possibility to perform high-throughput pH sensitivity analysis of protein sequences and structures.

A limitation of patchwork is that it does not capture pH-dependent structural dynamics that can occur upon pH shift. This can be achieved via molecular dynamics simulations at constant pH (43) comparing the results ob-

tained at different pH values. However, such approaches are computationally demanding and require expertise. On a more intuitive and accessible level, patchwork allows users to nonetheless predict potential structural rearrangements upon evaluation of gain or loss of noncovalent bonds caused by pH shift.

We believe patchwork will be an invaluable tool supporting research and teaching, deepening our mechanistic understanding of how pH impacts protein function.

DATA AVAILABILITY

The web server is freely available at <https://patchwork.biologie.uni-freiburg.de>.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We would like to thank the anonymous reviewers and members of the Di Ventura laboratory for their valuable feedback and suggestions.

FUNDING

German Ministry for Education and Research (BMBF) [031L0079 to B.D.V.]; Excellence Initiative of the German Federal and State Governments (BIOS Centre for Biological Signaling Studies) [EXC-294]; European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Framework Programme [101002044 to B.D.V.]. Funding for open access charge: European Research Council (ERC).

Conflict of interest statement. None declared.

REFERENCES

- Pace, C.N., Grimsley, G.R. and Scholtz, J.M. (2009) Protein ionizable groups: pK values and their contribution to protein stability and solubility. *J. Biol. Chem.*, **284**, 13285–13289.
- Schaefer, M., Sommer, M. and Karplus, M. (1997) pH-dependence of protein stability: absolute electrostatic free energy differences between conformations. *J. Phys. Chem. B*, **101**, 1663–1683.
- Mueller, E.A., Iken, A.G., Öztürk, M. A., Winkle, M., Schmitz, M., Vollmer, W., Di Ventura, B. and Levin, P.A. (2021) The active repertoire of *Escherichia coli* peptidoglycan amidases varies with physicochemical environment. *Mol. Microbiol.*, **116**, 311–328.
- Pathak, A.K. (2018) Effect of pH on the hinge region of influenza viral protein: a combined constant pH and well-tempered molecular dynamics study. *J. Phys. Condens. Matter*, **30**, 195101.
- Zhao, Y.J., He, W.M., Zhao, Z.Y., Li, W.H., Wang, Q.W., Hou, Y.N., Tan, Y. and Zhang, D. (2021) Acidic pH irreversibly activates the signaling enzyme SARM1. *FEBS J.*, **288**, 6783–6794.
- Schönichen, A., Webb, B.A., Jacobson, M.P. and Barber, D.L. (2013) Considering protonation as a posttranslational modification regulating protein structure and function. *Annu. Rev. Biophys.*, **42**, 289–314.
- Karshikoff, A. and Jelesarov, I. (2008) Salt bridges and conformational flexibility: effect on protein stability. *Biotechnol. Bioequip.*, **22**, 606–611.
- Yao, X., Chen, C., Wang, Y., Dong, S., Liu, Y.-J., Li, Y., Cui, Z., Gong, W., Perrett, S., Yao, L. *et al.* (2020) Discovery and mechanism of a pH-dependent dual-binding-site switch in the interaction of a pair of protein modules. *Sci. Adv.*, **6**, eabd7182.
- Hebditch, M. and Warwicker, J. (2019) Web-based display of protein surface and pH-dependent properties for assessing the developability of biotherapeutics. *Sci. Rep.*, **9**, 1969.
- Bartonek, L. and Zagrovic, B. (2019) VOLPES: an interactive web-based tool for visualizing and comparing physicochemical properties of biological sequences. *Nucleic Acids Res.*, **47**, W632–W635.
- Unni, S., Huang, Y., Hanson, R.M., Tobias, M., Krishnan, S., Li, W.W., Nielsen, J.E. and Baker, N.A. (2011) Web servers and services for electrostatics calculations with APBS and PDB2PQR. *J. Comput. Chem.*, **32**, 1488–1491.
- Cameselle, J.C., Ribeiro, J.M. and Sillero, A. (1986) Derivation and use of a formula to calculate the net charge of acid–base compounds. Its application to amino acids, proteins and nucleotides. *Biochem. Educ.*, **14**, 131–136.
- Isom, D.G., Castañeda, C.A., Cannon, B.R. and García-Moreno E.B. (2011) Large shifts in pK_a values of lysine residues buried inside a protein. *Proc. Natl Acad. Sci. U.S.A.*, **108**, 5260–5265.
- Alexov, E., Mehler, E.L., Baker, N., Baptista, A.M., Huang, Y., Milletti, F., Nielsen, J.E., Farrell, D., Carstensen, T., Olsson, M.H.M. *et al.* (2011) Progress in the prediction of pK_a values in proteins. *Proteins Struct. Funct. Bioinformatics*, **79**, 3260–3275.
- Dolinsky, T.J., Czodrowski, P., Li, H., Nielsen, J.E., Jensen, J.H., Klebe, G. and Baker, N.A. (2007) PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res.*, **35**, W522–W525.
- Olsson, M.H.M., Søndergaard, C.R., Rostkowski, M. and Jensen, J.H. (2011) PROPKA3: consistent treatment of internal and surface residues in empirical pK_a predictions. *J. Chem. Theory Comput.*, **7**, 525–537.
- Rostkowski, M., Olsson, M.H., Søndergaard, C.R. and Jensen, J.H. (2011) Graphical analysis of pH-dependent properties of proteins predicted using PROPKA. *BMC Struct. Biol.*, **11**, 6.
- Humphrey, W., Dalke, A. and Schulten, K. (1996) VMD: visual molecular dynamics. *J. Mol. Graph.*, **14**, 33–38.
- Piovesan, D., Minervini, G. and Tosatto, S.C.E. (2016) The RING 2.0 web server for high quality residue interaction networks. *Nucleic Acids Res.*, **44**, W367–W374.
- Jubb, H.C., Higuero, A.P., Ochoa-Montano, B., Pitt, W.R., Ascher, D.B. and Blundell, T.L. (2017) Arpeggio: a web server for calculating and visualising interatomic interactions in protein structures. *J. Mol. Biol.*, **429**, 365–371.
- Ferruz, N., Schmidt, S. and Höcker, B. (2021) ProteinTools: a toolkit to analyze protein structures. *Nucleic Acids Res.*, **49**, W559–W566.
- Abraham, M.J., Murtola, T., Schulz, R., Páll, S., Smith, J.C., Hess, B. and Lindahl, E. (2015) GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, **1–2**, 19–25.
- Barlow, D.J. and Thornton, J.M. (1983) Ion-pairs in proteins. *J. Mol. Biol.*, **168**, 867–885.
- Bhattacharyya, R., Samanta, U. and Chakrabarti, P. (2002) Aromatic–aromatic interactions in and around α -helices. *Protein Eng. Des. Sel.*, **15**, 91–100.
- Gallivan, J.P. and Dougherty, D.A. (1999) Cation– π interactions in structural biology. *Proc. Natl Acad. Sci. U.S.A.*, **96**, 9459–9464.
- Nakajima, K., Yokoyama, K., Koizumi, T., Koizumi, A., Asakura, T., Terada, T., Masuda, K., Ito, K., Shimizu-Ibuka, A., Misaka, T. *et al.* (2011) Identification and modulation of the key amino acid residue responsible for the pH sensitivity of neoculin, a taste-modifying protein. *PLoS One*, **6**, e19448.
- Ohkubo, T., Tamiya, M., Abe, K. and Ishiguro, M. (2015) Structural basis of pH dependence of neoculin, a sweet taste-modifying protein. *PLoS One*, **10**, e0126921.
- Chernova, M.N., Stewart, A.K., Barry, P.N., Jennings, M.L. and Alper, S.L. (2008) Mouse Ae1 E699Q mediates SO₄²⁻/anion_o exchange with [SO₄²⁻]-dependent reversal of wild-type pH_o sensitivity. *Am. J. Physiol. Cell Physiol.*, **295**, C302–C312.
- Sekler, I., Lo, R.S. and Kopito, R.R. (1995) A conserved glutamate is responsible for ion selectivity and pH dependence of the mammalian anion exchangers AE1 and AE2. *J. Biol. Chem.*, **270**, 28751–28758.
- Reimold, F.R., Stewart, A.K., Stolpe, K., Heneghan, J.F., Shmukler, B.E. and Alper, S.L. (2013) Substitution of transmembrane domain Cys residues alters pH_o-sensitive anion transport by AE2/SLC4A2 anion exchanger. *Pflüg. Arch. Eur. J. Physiol.*, **465**, 839–851.
- Jennings, M.L. and Smith, J.S. (1992) Anion–proton cotransport through the human red blood cell band 3 protein. Role of glutamate 681. *J. Biol. Chem.*, **267**, 13964–13971.
- Chernova, M.N., Jiang, L., Crest, M., Hand, M., Vandorpe, D.H., Strange, K. and Alper, S.L. (1997) Electrogenic sulfate/chloride exchange in *Xenopus* oocytes mediated by murine AE1 E699Q. *J. Gen. Physiol.*, **109**, 345–360.
- Bellotti, D., Rowińska-Żyrek, M. and Remelli, M. (2020) Novel insights into the metal binding ability of ZinT periplasmic protein from *Escherichia coli* and *Salmonella enterica*. *Dalton Trans.*, **49**, 9393–9403.
- Shirasuka, Y., Nakajima, K.-I., Asakura, T., Yamashita, H., Yamamoto, A., Hata, S., Nagata, S., Abo, M., Sorimachi, H. and Abe, K. (2004) Neoculin as a new taste-modifying protein occurring in the fruit of *Curculigo latifolia*. *Biosci. Biotechnol. Biochem.*, **68**, 1403–1407.
- Nakajima, K., Morita, Y., Koizumi, A., Asakura, T., Terada, T., Ito, K., Shimizu-Ibuka, A., Maruyama, J., Kitamoto, K., Misaka, T. *et al.* (2008) Acid-induced sweetness of neoculin is ascribed to its pH-dependent agonistic–antagonistic interaction with human sweet taste receptor. *FASEB J.*, **22**, 2323–2330.
- Shimizu-Ibuka, A., Morita, Y., Terada, T., Asakura, T., Nakajima, K., Iwata, S., Misaka, T., Sorimachi, H., Arai, S. and Abe, K. (2006) Crystal structure of neoculin: insights into its sweetness and taste-modifying activity. *J. Mol. Biol.*, **359**, 148–158.
- Romero, M.F., Fulton, C.M. and Boron, W.F. (2004) The SLC4 family of HCO₃⁻ transporters. *Mol. Aspects Med.*, **447**, 495–509.

38. Jennings, M.L. and Al-Rhaiyel, S. (1988) Modification of a carboxyl group that appears to cross the permeability barrier in the red blood cell anion transporter. *J. Gen. Physiol.*, **92**, 161–178.
39. Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G., Laydon, A. *et al.* (2022) AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.*, **50**, D439–D444.
40. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Židek, A., Potapenko, A. *et al.* (2021) Highly accurate protein structure prediction with AlphaFold. *Nature*, **596**, 583–589.
41. Zhekova, H.R., Pushkin, A., Kayık, G., Kao, L., Azimov, R., Abuladze, N., Kurtz, D., Damergi, M., Noskov, S.Y. and Kurtz, I. (2021) Identification of multiple substrate binding sites in SLC4 transporters in the outward-facing conformation: insights into the transport mechanism. *J. Biol. Chem.*, **296**, 100724.
42. Reithmeier, R.A.F., Casey, J.R., Kalli, A.C., Sansom, M.S.P., Alguet, Y. and Iwata, S. (2016) Band 3, the human red cell chloride/bicarbonate anion exchanger (AE1, SLC4A1), in a structural context. *Biochim. Biophys. Acta: Biomembr.*, **1858**, 1507–1532.
43. Socher, E. and Sticht, H. (2016) Mimicking titration experiments with MD simulations: a protocol for the investigation of pH-dependent effects on proteins. *Sci. Rep.*, **6**, 22523.