THE EMBO JOURNAL

A non-helical region in transmembrane helix 6 of hydrophobic amino acid transporter MhsT mediates substrate recognition.

Poul Nissen, Dorota Focht, Caroline Neumann, Joseph Lyons, Ander Eguskiza, Rikard Blunck, Lina Malinauskaite, Jonathan Javitch, Ilona Schwarz, and Matthias Quick
DOI: 10.15252/embj.2020105164

Corresponding author(s): Poul Nissen (pn@mbg.au.dk) , Matthias Quick (mq2102@cumc.columbia.edu)

Review Timeline:

Submission Date:	2nd Apr 20
Editorial Decision:	14th May 20
Revision Received:	3rd Jul 20
Editorial Decision:	5th Aug 20
Revision Received:	23rd Sep 20
Accepted:	1st Oct 20

Editor: Elisabetta Argenzio

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. Depending on transfer agreements, referee reports obtained elsewhere may or may not be included in this compilation. Referee reports are anonymous unless the Referee chooses to sign their reports.)

1st Editorial Decision 14th May 2020

Thank you for submitting your manuscript entitled "Role of transmembrane helix 6 in substrate recognition of the amino acid transporter MhsT" [EMBOJ-2020-105160] to The EMBO Journal. Please accept my apologies for the unusual length of the reviewing process due to a belated report. Your study has been sent to three reviewers for evaluation, whose reports are enclosed below.

As you can see, the referees consider the work potentially interesting. However, they also raise several criticisms that need to be addressed before they can support publication in The EMBO Journal.

Given the overall interest of your study, I am pleased to invite submission of a revised manuscript as indicated in the referee's reports. I would like to point it out that addressing all the referees' points in a conclusive manner, as well as a strong support from the referees, will be essential for publication in The EMBO Journal.

Please note that it is The EMBO Journal policy to allow only a single major round of revision and it is therefore important to resolve the main concerns at this stage.

When preparing your letter of response to the referees' comments, bear in mind that this will form part of the Review Process File and therefore will be available online to the community. For more details on our Transparent Editorial Process, please visit our website: http://emboj.embopress.org/about#Transparent_Process.

Before submitting your revision, primary datasets (and computer code, where appropriate) produced in this study need to be deposited in an appropriate public database (see http://msb.embopress.org/authorguide#dataavailability). Remember to provide a reviewer password if the datasets are not yet public.

We usually expect to receive revised manuscripts within three months of the first decision. We are aware that many laboratories cannot function at full capacity during the current COVID-19/SARS-CoV-2 pandemic and may relax this deadline. Also, we can extend our 'scooping protection policy' to cover the period required for a full revision to address all of the referees' points. Please inform us as soon as a paper with related content published elsewhere.

Thank you again for the opportunity to consider this work for publication, and please feel free to contact me with any questions about submission of the revised manuscript to The EMBO Journal.

I look forward to your revision.

Referee #1:

The manuscript by Focht and co-workers presents an interesting study about the structural basis for substrate selectivity of the amino acid transporter MhsT. They provide very convincing data demonstrating that the unwound portion of TM6 provides a flexible binding pocket for substrates

and that the key residue in this motif M236 can adopt different conformations depending on the substrate being transported. It is well established that this region of MhsT, and also all the other members of the SLC6 family of transporters, plays a key role in determining selectivity of the side chain of the transported substrates. In fact, in the original papers on the crystal structures of LeuT (Yamishita et al., 2005 and Singh et al 2008) it was suggested that the residue that corresponds to M236 in LeuT, F259 is a major determinant of the volume of side chains of substrates that may fit into the binding site. It was pointed out that the corresponding residue in the GlyTs is a W residue and subsequent studies demonstrated that the side chain of this residue is a major determinant of the volume of substrates that can be accommodated in the site. So, whilst the identification of the role of this region is not novel, the new contribution provided by this study is that the conformation of the M236 side chain, together with the side chain of E66, determines how the side chains of both aromatic and aliphatic amino acids can fit into the site. This novel observation is clearly demonstrated.
I have no major concerns of a technical nature relating to the crystallography or of the clarity of the work presented. I do feel that there is a bit of redundancy in the figures presented. It appears that the same data has been presented in multiple forms showing subtly different aspects of the binding sites.
The legend for the supplementary data presented in Supp Fig 9 does not relate to the data presented. The data in the first graph of part A of S9 has considerable error and this has the potential to generate considerable error in the rates of uptake obtained. Either the figure legend should be rewritten to reflect the data presented or is another figure missing?

Referee #2:

Focht et al. report a series of generally high resolution (5/6) structures of the bacterial SLC6 homolog MhsT bound to three aromatic (Tyr, 4-fluoro-Phe, Phe) and three aliphatic (Ile, Leu, and Val) substrates. All structures are in the inward-facing occluded state. The main finding is that the 'GMG' loop follows the substrate in a way such that the volume of the substrate cavity expands or shrinks to accommodate the size of the substrate. The authors claim test whether the 'M' of the loop is key to substrate selectivity. A single change M□F is able to eliminate the transport of aromatic amino acids but preserves the ability to transport aliphatic ones. Although the adaptation of the GMG loop to the substrate volume is not particularly surprising, the selectivity change is a nice demonstration that the authors understand something about substrate selection from these studies. A slightly more comprehensive mutational analysis would likely be informative, but may be outside the realm of possibility in the current era of closed, or highly restricted laboratories. I imagine that the selectivity change comes from the increased volume and rigidity of the F vs. M. Commenting on this point, and how other substitutions might influence substrate selection would help frame a larger context for this work.

There are typos throughout the Figure legends. Descriptions are also missing. Ex. in Fig. 3C, which structure contributes the green loops?

Referee #3:

The Nissen and Quick groups present a series of crystal structures of MhsT, a prokaryotic member

of the SLC6 family, and reveal conformational adaptations in a membrane-embedded region to accommodate differently sized amino acid substrates. They also perform in vivo transport assays to confirm the importance of a methionine residue in this region for substrate selectivity.

Overall quality of the work is excellent. Although I would have liked to see a bit more in depth functional/dynamics analysis, the observations are insightful. The manuscript is also well written, although the discussion could be shortened considerably.

I have a few specific points:
1. Figure 1, legend: line 464: "hydrophobic amino acids". Also non-hydrophobic ones are shown
2. Figure S9 is not discussed much in the text and it is not clear what tetra, penta represent, and why these (mutants?) were studied
3. Glu66: the authors discuss this residue to great length. Why have they not simply mutated it to test their interpretation based on the structures? I suggest to either do the experiments, or remove much of the speculative discussion.
4. Line 297: Other residues were tested: show data
5. Line 298: comparable levels: show data
6. Line 219: "that may offset" I do not understand what the authors want to say. Why is offset needed?
7. Line 168: "induced fit". It is not possible to distinguish induced fit from conformational selection just by looking at structures (see for instance https://pubmed.ncbi.nlm.nih.gov/23502425/ This statement screams for smFRET experiments (which the authors are capable of doing!)
8. Line 91: "this modulation" not clear what it refers to. Modulation has not been mentioned.
9. Line 92: check grammar; sentence does not flow
10. Line 77: close ortolog. "close" is misplaced here.

EMBOJ-2020-105164 – Letter to referees

The figures have been rearranged according to the referees' comments:

Main text figures:

Figure 1

Figure 2 (previous figure 4)

Figure 3

Figure 4 (previous figure 5)

Figure 5 (previous figure 6)

Expanded View figures:

Figure EV1 (previous figure S6)

Figure EV2 (previous figure S7)

Figure EV3 (previous figure S8)

Figure EV4 (previous figure S9)

Appendix figures:

Appendix Figure S1 (previous figure 2)

Appendix Figure S2 (previous figure S1)

Appendix Figure S3

Appendix Figure S4

Appendix Figure S5 (previous figure S2)

Appendix Figure S6 (previous figure S6)

Referee #1

1. *I do feel that there is a bit of redundancy in the figures presented. It appears that the same data has been presented in multiple forms showing subtly different aspects of the binding sites*

We acknowledge this point and have moved figure 2 to the appendix (now called Appendix Figure S1), and old figure 4 into the place of a new figure 2.

2. *The legend for the supplementary data presented in Supp Fig 9 does not relate to the data presented. –*

We very much apologize for mistakes in the submission process and internal communications - the figure S9 has been exchanged with the correct, originally

intended figure (the figure is now moved to the Expanded View figures and is called Figure EV4).

3. *The data in the first graph of part A of S9 has considerable error and this has the potential to generate considerable error in the rates of uptake obtained. Either the figure legend should be rewritten to reflect the data presented or is another figure missing?*

The figure has been exchanged with the correct one, we apologize (now called Figure EV4).

Referee #2

1. *A slightly more comprehensive mutational analysis would likely be informative, but may be outside the realm of possibility in the current era of closed, or highly restricted laboratories. –*

In fact, more comprehensive mutagenesis was done earlier, data shown in corrected Fig S9 (now called Figure EV4). We sincerely apologize for the confusion.

2. *I imagine that the selectivity change comes from the increased volume and rigidity of the F vs. M. Commenting on this point, and how other substitutions might influence substrate selection would help frame a larger context for this work.*

We believe that with the more extensive mutations available now to the reader, the comparison to sequence motifs of other SLC6 transporters, and the discussion points included in that regard, we will cover the salient points of the F vs. M.

3. *There are typos throughout the Figure legends. Descriptions are also missing. Ex. in Fig. 3C, which structure contributes the green loops?*

We have corrected typos and added missing descriptions.

Referee #3

…..The discussion could be shortened considerably.

We have moved material better suited for the Results sections and streamlined the discussion points and removed as much as possible redundant and unfocused points.

1. *Figure 1, legend: line 464: "hydrophobic amino acids". Also non-hydrophobic ones are shown –*

The sentence has been corrected

2. *Figure S9 is not discussed much in the text and it is not clear what tetra, penta represent, and why these (mutants?) were studied*

The figure has been exchanged to the correct one, and referred to in appropriate parts of the manuscript; triple18 is explained in the legend (now called Figure EV4).

3. *Glu66: the authors discuss this residue to great length. Why have they not simply mutated it to test their interpretation based on the structures? I suggest to either do the experiments, or remove much of the speculative discussion*

The residue has been extensively studied in other transporters, and the MhsT structures offer a good opportunity to revisit the structural role of this residue, comparing outward- and inward-oriented substrate complexes of LeuT and MhsT, respectively. If more advanced assays will later be developed for MhsT dynamics (e.g. smFRET, see further below) Glu66 would seem like a very attractive target for mutational studies in MhsT.

4. *Line 297: Other residues were tested: show data –*

Data were to be shown in Fig S9, now exchanged to the correct figure and called Figure EV4.

5. *Line 298: comparable levels: show data*

same as above

6. Line 219: "that may offset" I do not understand what the authors want to say. Why is offset needed?

We have changed this sentence
ORIGINAL: ***Interestingly, in both SLC6A18 and SLC6A19 the AFG motif precedes an additional glycine that may *offset* this reduced flexibility and bulkier central residue..***

We have changed it to the following: ***"Interestingly, in both SLC6A18 and SLC6A19 the AFG motif is followed by an additional glycine that may provide additional, spatial flexibility"***

7. *Line 168: "induced fit".It is not possible to distinguish induced fit from conformational selection just by looking at structures (see for instance https://pubmed.ncbi.nlm.nih.gov/23502425/ This statement screams for smFRET experiments (which the authors are capable of doing!) –*

We agree with the Reviewer that smFRET measurements would be a fantastic addition to provide further insight into the mechanistic interpretation of our data. However, while we reported on a smFRET-based assay for imaging MhsT-mediated transport in MhsT-incorporated proteoliposomes that contain the smFRET sensor (the fluorophore-labeled LIV binding protein) (Fitzgerald et al., 2019), establishing smFRET-based dynamics measurements within MhsT, as the Reviewer suggests, is

far from trivial. In fact, we have tried mutating positions in MhsT that align with those successfully targeted in LeuT for measuring intra- and extracellular ligand-induced distance changes (Terry et al., 2018), but the mutations led to reduced expression and stability of these MhsT constructs, so there is no simple adoption of methods that were established for LeuT over several years. Given the current restrictions caused by the COVID-19 pandemic, it is uncertain when these studies could even commence. We feel that, in light of the overall positive comments by the Reviewers, the publication of this manuscript should not be delayed due to the lack of data that would require complex and time-consuming technology development to achieve. Should we establish a new experimental platform it would warrant an independent publication we think.

8. *Line 91: "this modulation" not clear what it refers to. Modulation has not been mentioned.*
9. *Line 92: check grammar; sentence does not flow –*
10. *Line 77: close ortolog. "close" is misplaced here.*

These correction have been made.

1st Revision - Editorial Decision 5th Aug 2020

Thank you for submitting your revised manuscript. Your study has been re-reviewed by the original referees and we have now received their reports, which are enclosed below.

As you will see, the referees find that their criticisms have been sufficiently addressed and recommend the manuscript for publication, pending some textual changes.

In addition, there are a few editorial issues concerning text and figures that I need you to address before we can officially accept your manuscript.

Referee #1:

The concerns raised have been adequately addressed

Referee #2:

The authors have addressed the previous comments. The work is well done and clearly presented.

Referee #3:

The authors have dealt with most concerns adequately, with one exception: the discussion of the induced fit speculation. Although I fully understand that experimental testing by smFRET is not possible, the authors must make clear that they do not have experimental evidence to distinguish between induced fit and conformational selection. In fact they should explain why they dismiss the possibility of conformational selection without mention. This can all be done textually, but it MUST be done to avoid misleading ignorant readers.

2nd Authors' Response to Reviewers 28th Sep 2020

It is our great pleasure to submit the adustments to our revised manuscript “Role of a non-helical region of transmembrane helix 6 in the substrate recognition mechanism of the hydrophobic amino acid transporter MhsT” for publication as a Research Article in EMBO Journal. We fully concur to reviewer 3’s comment on induced fit (a classical term, but indeed not necessarily fulfilling) vs. conformational selection and have changed the sentence in the following way:

Overall changes in the volume of the MhsT hydrophobic cavity upon binding of different substrates could follow an “induced-fit” mechanism (Klingenberg, 2005; Nyola et al, 2010) or conformational selection (LeVine & Weinstein, 2014; Hammes et al. 2009). Specifically the hydrophobic nature of the substrate and its binding pocket disfavors the possibility of increased solvation compensating for the deficit in substrate volume, thus promoting an intrinsic structural fit of the substrate binding site – a mechanism, which is also reminiscent of the movement of the unwound segment of TM1 to compensate the empty hydrophobic-lined substrate binding site in the occluded return state of LeuT and other SLC6 family transporters (Malinauskaite et al, 2016).

2nd Revision - Editorial Decision 1st Oct 2020

I am pleased to inform you that your manuscript has been accepted for publication in The EMBO Journal.

Congratulations!

EMBO PRESS

YOU MUST COMPLETE ALL CELLS WITH A PINK BACKGROUND ↓

PLEASE NOTE THAT THIS CHECKLIST WILL BE PUBLISHED ALONGSIDE YOUR PAPER

Corresponding Author Name: Poul Nissen
Journal Submitted to: EMBO Journal
Manuscript Number: EMBOJ-2020-105164

Reporting Checklist For Life Sciences Articles (Rev. June 2017)

This checklist is used to ensure good reporting standards and to improve the reproducibility of published results. These guidelines are consistent with the Principles and Guidelines for Reporting Preclinical Research issued by the NIH in 2014. Please follow the journal's authorship guidelines in preparing your manuscript.

A- Figures

1. Data

The data shown in figures should satisfy the following conditions:

- ➔ the data were obtained and processed according to the field's best practice and are presented to reflect the results of the experiments in an accurate and unbiased manner.
- ➔ figure panels include only data points, measurements or observations that can be compared to each other in a scientifically meaningful way.
- ➔ graphs include clearly labeled error bars for independent experiments and sample sizes. Unless justified, error bars should not be shown for technical replicates.
- ➔ if n< 5, the individual data points from each experiment should be plotted and any statistical test employed should be iustified
- ➔ Source Data should be included to report the data underlying graphs. Please follow the guidelines set out in the author ship guidelines on Data Presentation.

2. Captions

Each figure caption should contain the following information, for each panel where they are relevant:

- ➔ a specification of the experimental system investigated (eg cell line, species name).
- ➔ the assay(s) and method(s) used to carry out the reported observations and measurements
- ➔ an explicit mention of the biological and chemical entity(ies) that are being measured.
- ➔ an explicit mention of the biological and chemical entity(ies) that are altered/varied/perturbed in a controlled manner.

- ➔ the exact sample size (n) for each experimental group/condition, given as a number, not a range;
- ➔ a description of the sample collection allowing the reader to understand whether the samples represent technical or biological replicates (including how many animals, litters, cultures, etc.).
- ➔ a statement of how many times the experiment shown was independently replicated in the laboratory.
- ➔ definitions of statistical methods and measures:
  - common tests, such as t-test (please specify whether paired vs. unpaired), simple $\chi 2$ tests, Wilcoxon and Mann-Whitney tests, can be unambiguously identified by name only, but more complex techniques should be described in the methods section;
  - are tests one-sided or two-sided?
  - are there adjustments for multiple comparisons?
  - exact statistical test results, e.g., P values = x but not P values < x;
  - definition of 'center values' as median or average;
  - definition of error bars as s.d. or s.e.m.

Any descriptions too long for the figure legend should be included in the methods section and/or with the source data.

In the pink boxes below, please ensure that the answers to the following questions are reported in the manuscript itself. Every question should be answered. If the question is not relevant to your research, please write NA (non applicable). We encourage you to include a specific subsection in the methods section for statistics, reagents, animal models and human subjects.

B- Statistics and general methods

Please fill out these boxes ↓ (Do not worry if you cannot see all your text once you press return)

1.a. How was the sample size chosen to ensure adequate power to detect a pre-specified effect size?	Per standard practice in the lab, pilot experiments were performed to identify requirements to accurately assess the test parameters. Following, transport experiments were performed a minimum of 5 times using different batches of cells expressing the appropriate transporter construct (and control cells). Each experiment was performed as technical duplicates or triplicates.
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	n.a.
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	n.a.
3. Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. randomization procedure)? If yes, please describe.	n.a.
For animal studies, include a statement about randomization even if no randomization was used.	n.a.
4.a. Were any steps taken to minimize the effects of subjective bias during group allocation or/and when assessing results (e.g. blinding of the investigator)? If yes please describe.	n.a.
4.b. For animal studies, include a statement about blinding even if no blinding was done	n.a.
5. For every figure, are statistical tests justified as appropriate?	n.a.
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	n.a.

USEFUL LINKS FOR COMPLETING THIS FORM

http://www.antibodypedia.com
http://1degreebio.org
http://www.equator-network.org/reporting-guidelines/improving-bioscience-research-repor

http://grants.nih.gov/grants/olaw/olaw.htm
http://www.mrc.ac.uk/Ourresearch/Ethicsresearchguidance/Useofanimals/index.htm
http://ClinicalTrials.gov
http://www.consort-statement.org
http://www.consort-statement.org/checklists/view/32-consort/66-title

http://www.equator-network.org/reporting-guidelines/reporting-recommendations-for-tumo

http://datadryad.org

http://figshare.com

http://www.ncbi.nlm.nih.gov/gap

http://www.ebi.ac.uk/ega

http://biomodels.net/

http://biomodels.net/miriam/
http://jjj.biochem.sun.ac.za
http://oba.od.nih.gov/biosecurity/biosecurity_documents.html
http://www.selectagents.gov/

Is there an estimate of variation within each group of data?	n.a.
Is the variance similar between the groups that are being statistically compared?	n.a.

C- Reagents

6. To show that antibodies were profiled for use in the system under study (assay and species), provide a citation, catalog number and/or clone number, supplementary information or reference to an antibody validation profile. e.g., Antibodypedia (see link list at top right), 1DegreeBio (see link list at top right).	n.a.
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	n.a.

* for all hyperlinks, please see the table at the top right of the document

D- Animal Models

8. Report species, strain, gender, age of animals and genetic modification status where applicable. Please detail housing and husbandry conditions and the source of animals.	n.a.
9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.	n.a.
10. We recommend consulting the ARRIVE guidelines (see link list at top right) (PLoS Biol. 8(6), e1000412, 2010) to ensure that other relevant aspects of animal studies are adequately reported. See author guidelines, under 'Reporting Guidelines'. See also: NIH (see link list at top right) and MRC (see link list at top right) recommendations. Please confirm compliance.	n.a.

E- Human Subjects

11. Identify the committee(s) approving the study protocol.	n.a.
12. Include a statement confirming that informed consent was obtained from all subjects and that the experiments conformed to the principles set out in the WMA Declaration of Helsinki and the Department of Health and Human Services Belmont Report.	n.a.
13. For publication of patient photos, include a statement confirming that consent to publish was obtained.	n.a.
14. Report any restrictions on the availability (and/or on the use) of human data or samples.	n.a.
15. Report the clinical trial registration number (at ClinicalTrials.gov or equivalent), where applicable.	n.a.
16. For phase II and III randomized controlled trials, please refer to the CONSORT flow diagram (see link list at top right) and submit the CONSORT checklist (see link list at top right) with your submission. See author guidelines, under 'Reporting Guidelines'. Please confirm you have submitted this list.	n.a.
17. For tumor marker prognostic studies, we recommend that you follow the REMARK reporting guidelines (see link list at top right). See author guidelines, under 'Reporting Guidelines'. Please confirm you have followed these guidelines.	n.a.

F- Data Accessibility

18: Provide a "Data Availability" section at the end of the Materials & Methods, listing the accession codes for data generated in this study and deposited in a public database (e.g. RNA-Seq data: Gene Expression Omnibus GSE39462, Proteomics data: PRIDE PXD000208 etc.) Please refer to our author guidelines for 'Data Deposition'. Data deposition in a public repository is mandatory for: a. Protein, DNA and RNA sequences b. Macromolecular structures c. Crystallographic data for small molecules d. Functional genomics data e. Proteomics and molecular interactions	PDB structures deposited and accepted; PDB IDs: 6YU2(MhsT-L-isoleucine), 6YU3 (L-phenylalanine), 6YU4 (MhsT-L-4F-phenylalanine), 6YU5 (MhsT-L-valine), 6YU6 (L-leucine), 6YU7 (L-tyrosine)
19. Deposition is strongly recommended for any datasets that are central and integral to the study; please consider the journal's data policy. If no structured public repository exists for a given data type, we encourage the provision of datasets in the manuscript as a Supplementary Document (see author guidelines under 'Expanded View' or in unstructured repositories such as Dryad (see link list at top right) or Figshare (see link list at top right).	PDB structures and mtz files will be available on https://www.rcsb.org
20. Access to human clinical and genomic datasets should be provided with as few restrictions as possible while respecting ethical obligations to the patients and relevant medical and legal issues. If practically possible and compatible with the individual consent agreement used in the study, such data should be deposited in one of the major public access-controlled repositories such as dbGAP (see link list at top right) or EGA (see link list at top right).	n.a.
21. Computational models that are central and integral to a study should be shared without restrictions and provided in a machine-readable form. The relevant accession numbers or links should be provided. When possible, standardized format (SBML, CellML) should be used instead of scripts (e.g. MATLAB). Authors are strongly encouraged to follow the MIRIAM guidelines (see link list at top right) and deposit their model in a public database such as Biomodels (see link list at top right) or JWS Online (see link list at top right). If computer source code is provided with the paper, it should be deposited in a public repository or included in supplementary information.	n.a.

G- Dual use research of concern

22. Could your study fall under dual use research restrictions? Please check biosecurity documents (see link list at top right) and list of select agents and toxins (APHIS/CDC) (see link list at top right). According to our biosecurity guidelines, provide a statement only if it could.	no