E-Rare final summarizing report of the ERAAT project consortium to EC, 25th July 2022

Public summary (can be disclosed)

Major results

Within the ERAAT project, a primary assay suitable for high throughput screening (HTS) of small molecules was developed. HTS was conducted and more than 570.000 compounds were tested. About 800 primary hits were identified (IRE1 inducers) that were subsequently characterized in secondary assays, chemically clustered and analyzed by cheminformatics. In addition, an *in silico* exploration of the IRE1 kinase was conducted with a set of analogs of IRE1 oligopeptide sequences exhibiting inhibitory activities. The pharmacophore model, together with the binding hypotheses resulting from this screen, were then utilized as filters in an *in silico* screening campaign and potential hit candidates were identified and characterized. Molecular modelling was used to explore the possible modes of action of the IRE1 activator compounds that resulted in hit molecules which were active in biological experiments.

With this project, the hypothesis that IRE1 activators increase the secretion of ZA1AT could be de-validated. Based on our results, it can be assumed that IRE1-activation is generally not sufficient to significantly catalyze the export of correctly or incorrectly folded proteins from the ER or autophagy.

Scientific report

State of the art

The ER regulates many cellular processes and its dysfunction underlies many diseases resulting from imbalanced cellular homeostasis, also termed proteostasis. A large number of diseases results from the failure of a protein to exit the ER. As a result mutant, but otherwise (fully) functional proteins such as ZA1AT become retained in the ER where it aggregates. In the case of soluble lumenal proteins, this includes ATD, leading to lung disease and, in some cases, to liver disease. As a member of the serine protease inhibitor (serpin) family, A1AT is the most abundant antiprotease in the serum. This protein is mainly produced by hepatocytes in the liver, but also in yet unknown amounts by other cellular tissues such as lung alveolar cells lining the interstitial space and by lung macrophages. The adult liver secretes more than one gram of A1AT per day into the bloodstream, where it functions as a general inhibitor of serine proteases, including neutrophil elastase, cathepsin G and proteinase-3. In the lung, its activity is critical for the inhibition of neutrophil elastase, a key component of the extracellular matrix that is responsible for the elastic properties of the lung. Therefore, hepatic production of A1AT is necessary for maintaining the protease/anti-protease balance in the serum. ATD is a genetic disorder, which is characterized by low serum levels of A1AT, the main protease inhibitor (PI) in human serum. Consequently, serum A1AT–deficient patients have an increased risk of developing severe lung disease and, depending on the particular A1AT variant, liver disease. Indeed, the general hallmark of all forms of A1AT disease is panacinar emphysema as early as the third to fifth decade of life, reflecting reduction of serum pools of A1AT. Patients homozygous for a Glu342 to Lys substitution found in the Z variant have a folding defect in A1AT that leads not only to inefficient export from the ER of the soluble monomer pool, but also to the accumulation of protein aggregates in the ER that form through loop-sheet polymerization of soluble monomers. Hepatic cells partially protect themselves from these aggregates using autophagic pathways that clear aggregates of polymers from the ER. When clearance is inadequate, accumulated aggregates trigger stress-signalling pathways that contribute to further liver dysfunction, including neonatal hepatitis, juvenile cirrhosis, and hepatocellular carcinoma. It is therefore essential to
develop therapies for these pathologies to replace orthotopic liver transplantation that thus far is the only option to treat the liver damages caused by ATD and that replacement therapies for A1AT have not proven effective yet.

Non-public section (confidential, due to IP protection reasons it should not be disclosed)

Aims and scientific approach

WP1 Compound logistics, preparation of compounds on assay ready plates (ARPs).
MS1 Provision of screening compounds for HTS.
WP2 Assays – Set up of all required assays for primary screen and hit identification as well as a set of secondary assays for hit profiling.
1. Transfer of primary IRE1-XBP1s splicing assay from Inserm to LDC and adaptation for HTS.
2. Biochemical IRE1 activation assay for hit confirmation.
3. RIDD assay to identify XBP1 activators and RIDD inhibitors.
4. ATZ secretion assay in a liver cell line, e.g. Huh-7 cells; counter assay for testing secretion of a wild type secreted protein.
5. Cellular assays/biological validation: Testing of initial hit compounds in cell lines to assess toxicity and cellular interference with UPR signaling.
6. Additional ER-stress reporters & counter screens for further characterization of hits.
7. ADMET assays: Pharmacological assays for assessment of pharmacological properties of the compounds.
MS2 Primary assay for HTS and secondary assays developed.
WP3 HTS & hit verification: HTS: Identification of ATZ secretion inducers by screening of diverse compound libraries. Hit verification: Verification of all primary hits in dose range experiments to determine their AC50-values. Summary of confirmed hit compounds that show reproducible activity in a primary hit list.
MS3 Primary hit list (AC50-values determined, clustered as chemical classes).
WP4 Hit validation - secondary assays: Validation of all confirmed hits of the primary hit list in a cascade of secondary assays developed in WP2. Clustering of active hits in different compound classes.
MS4 Validated hit list.
WP5 Peptidomimetics – Co-crystallization of the FIRE peptide with IRE1. Solving of crystal structures to elucidate the binding mode of the FIRE peptide; chemistry-based peptidomimetics approach to back up a potential shortage of primary hits from HTS. Development of a structure-based pharmacophore model. Identification of small substances which mimic the FIRE peptide binding site. Testing pf peptidomimetics in the established assays and including in the validated hit list.
MS5 Peptidometics identified.
WP6 Co-crystallization of validated hits from HTS with IRE1 – Use of a validated IRE1 activator for cocrystallization approaches with IRE1.
MS6 IRE1/IRE1 modulator co-crystals.
WP7 Hit-to-Lead (H2L) optimization - Prioritized validated hits as well as potential IRE1 activators from the peptidomimetic approach serve as chemical starting points for medicinal chemistry-based optimization. Improvement of quality of hit compounds through repeated cycles of organic synthesis and testing of all new compounds in the established assay cascade.
MS7 Lead candidate selected.
WP8 DMPK/Tox studies - *In vivo* pharmacokinetic (PK) and tolerability studies (non-GLP level, standard mouse and rat strains).
MS8 DMPK/Tox results acceptable. Testing of disease reduction in a therapeutic animal model for ATD (PiZ mouse model).
MS9 Lead series with proven efficacy in animal model.

Results/Achievements

Describe the activities that have been accomplished in order to fulfil the aims described in the network proposal. Specify the main results obtained and their relevance. Estimate the current degree of completion of the planned objectives. Please structure the report according to the aims. Describe

also the collaboration with the participating partners. Refer explicitly to common milestones and deliverables achieved during the year. Describe sharing of facilities/resources within the consortium.

Within the ERAAT project, a primary assay (primary IRE1-XBP1 splicing assay) for HTS of small molecule compound libraries was developed and 560.000 compounds were screened (WP2). Several secondary assays were developed at LDC and in the partners´ labs such as a biochemical IRE1 kinase assay, ER-stress reporter assays, protein trafficking assay, pharmacological ADMET assays, ZA1AT secretion assay, ZA1AT aggregation assay (WP2, MS2 reached), and others (see workplan above). Screening compounds were provided for HTS (WP1, MS1 reached). HTS was conducted (WP3) and about 800 primary hits (IRE1 inducers) were identified. These primary hits were summarized in a primary hitlist (MS3 reached). The primary hits were analyzed and characterized in several assays (WP4) and chemically clustered (MS4 reached).
In addition to the HTS approach, an *in silico* exploration of the IRE1 kinase was conducted with a set of analogs of IRE1 oligopeptide sequences exhibiting inhibitory activities. The pharmacophore model, together with the binding hypotheses resulting from this screen, was then utilized as filter in an *in silico* screening campaign and potential hit candidates were identified and characterized.

Molecular modelling was used to explore the possible modes of action of the IRE1 activator compounds. In summary, 35 virtual hits were identified by docking compounds from a database to the active site of IRE1. These compounds were tested against human IRE1 in a cell free RNase activity assay. Their binding modes of confirmed hits were analyzed by induced fit docking and the best models were subjected to molecular dynamics (MD) simulations. These calculations revealed their molecular mechanism of action. The activity of some hit molecules were confirmed in biological experiments. However, an obvious limitation of these compounds is their mediocre activity that should be optimized before evaluating them in more sophisticated cellular disease biology models.

Interestingly, all primary hit compounds activated IRE1, but only one fully verified activator hit from 273 confirmed hits enhanced the secretion of the ZA1AT mutant in the various cell-based (Z)A1AT secretiona assays to a low level. After testing in a set of different assays, we were facing the unexpected situation of having very few ZA1AT compound correctors that are not dependent on the increased activity of IRE1, i.e. there was no correlation between activation of IRE1 and induction of cellular (Z)A1AT secretion. Therefore, we decided to investigate this off-target effect of these few IRE1 activitators to learn more about biologically active compounds that allow correction of ZA1AT polymerization or increased secretion. To do so, we chose a platform of transcriptional ER-stress reporters to identify and investigate these steps. This so-called IPPER sentinel reporter platform ('Integrated Platform for the Pharmacology of the ER') contains a comprehensive set of beacon reporters to monitor common cellular changes. The reporters focus on but are not limited to monitoring ER stress, quality control steps, protein folding & degradation, and the unfolded protein response with its three UPR sensors ATF6, PERK, and IRE1. These luciferase-based reporters consist of promoters and transcription factor response elements, thereby improving detection sensitivity for proteostatic and other pathways. In contrast to e.g. RNAseq's overwhelming data flood, the IPPER platform overcomes the 'cocktail-party problem' by focusing on specific sentinel reporters. IPPER performance was benchmarked using quality chemical probes in kidney, lung, and liver cells. IPPER, combined with standard immunoblotting for key signaling components, revealed that in contrast to the non-corrector IRE1 activator hits of the Z-mutant, the correcting Ire1 activators have side or off-target effects that may explain their correcting abilities. The side effects resemble partial activation of the integrated stress response and autophagy.

The estimated degree of completion of the project is about 60%.

The collaboration and exchange of materials between the partners LDC (Germany), INSERM (France), McGill University (Canada), and Semmelweis University (Hungary) was excellent. Compounds and

plates were shipped on a regular basis for testing and profiling in the different partners´ labs. Regular meetings were held where data were exchanged and discussed. Due to Covid-19, most of the project meeting were organized virtually.

Discussion

Describe the implications of your results. Describe also problems that have hindered the achievement of the planned objectives (if any) and/or alternative work plans and/or changes with respect to the network proposal.

In comparison to the project plan, not all milestones could be reached due to the following reasons: (1) during the development of the primary ZA1AT secretion assay transient transfection did not work and a stable cell line needed to be produced, (2) technical issues had to be solved in compound logistics and transfer to the partners, (3) devalidating the published hypothesis of IRE1 activators leading to increased secretion of (Z)A1AT, (4) Covid-19 caused many obstacles (restricted access to labs, severe difficulties in aquirement of consumables). Issues 1 & 2 were solved with some time delays, issue 3 was thoroughly investigated and therefore a large part of the scheduled project development time, whereas (4), the Covid-19 related difficulties led to a significant delay in the project. As a consequence, WPs 5,6,7,8 could not be performed as planned and the respective milestones were not completely met. However, the project results provide a very good basis for further but independent investigations of ATD as well as on IRE1 activation, such as testing either cell-active hit compounds with off-target activities or peptides for ZA1AT secretion on disease-relevant iPSCs derived from homozygous ZA1AT patient primary cells. In summary and quite surprisingly, with this work we were able to devalidate the published hypothesis of IRE1 activation leading to increased (Z)A1AT secretion or inhibition of A1AT polymerization.

Semmelweis

No publications, one paper is planned about molecular modelling.

Two PhD students in the project (a female and a male student)

LDC

No publications, press release: https://www.lead-discovery.de/wp-content/uploads/2020/01/Press-release-ERAAT-final_EC.pdf

No students involved

McGill

no publications

no students involved

Inserm

Publications:

Pelizzari-Raymundo D, Doultsinos D, Pineau R, Sauzay C, Koutsandreas T, Carlesso A, Gkotsi E, Negroni L, Avril T, Guillory X, Eriksson LA, Chevet E. A novel blood brain barrier-permeable IRE1 kinase inhibitor sensitizes glioblastoma to chemotherapy in mice. **chemRxiv** DOI: 10.26434/chemrxiv-2022-2ld35

Amarasinghe KN, Pelizzari-Raymundo D, Carlesso A, Chevet E, Eriksson LA, Jalil Mahdizadeh S. Sensor dimer disruption as a new mode of action to block the IRE1-mediated unfolded protein response. **Comput Struct Biotechnol J.** 2022 Mar 29;20:1584-1592.

Papaioannou A, Centonze F, Metais A, Maurel M, Negroni L, Gonzalez-Quiroz M, Mahdizadeh SJ, Svensson G, Zare E, Blondel A, Koong AC, Hetz C, Pedeux R, Tremblay ML, Eriksson LA, Chevet E. Stress-induced tyrosine phosphorylation of RtcB modulates IRE1 activity and signaling outputs. **Life Sci Alliance.** 2022 Feb 22;5(5):e202201379.

Langlais T, Pelizzari-Raymundo D, Mahdizadeh SJ, Gouault N, Carreaux F, CHEVET E, Eriksson LA, Guillory X. Structural and molecular bases to IRE1 activity modulation. **Biochem J**. 2021 Aug 13;478(15):2953-2975. Review

Doultsinos D, Carlesso A, Chintha C, Paton JC, Paton AW, Samali A, CHEVET E, Eriksson LA. Peptidomimetic-based identification of FDA-approved compounds inhibiting IRE1 activity. **FEBS J.** 2021 Feb;288(3):945-960.

Raymundo DP, Doultsinos D, Guillory X, Carlesso A, Eriksson LA, CHEVET E. Pharmacological Targeting of IRE1 in Cancer. **Trends Cancer.** 2020 Dec;6(12):1018-1030. Review

Patent

Submission number: 1000505341

Application number: EP22305014.7

File No.to be used for priority declarations: EP22305014

Date of receipt: 07January2022

Applicant: INSERM (Institut National de la Santé et de la Recherche Médicale)

Inventeurs: E Chevet, LA Eriksson, X Guillory, D Pelizzari, D Doultsinos, T Langlais, S Sueron

COMPOUNDS CONTAINING A HYDROXYPHENYL MOIETY AND THEIR USE - The present invention relates to urea, oxalamide, amide, thiourea, carbamate or ester compounds, in particular urea compounds, containing a hydroxyphenyl or phenyl moiety, in particular a hydroxyphenyl moiety, including their pharmaceutically acceptable salts and solvates which are useful as sensitizers for chemotherapy of cancer cells, particularly in glioblastoma, and are useful as therapeutic compounds, particularly in the treatment of cancers that may be treated by alkylating agents, such as temozolomide.

Summary of Semmelweis University work in ERAAT IRE1 project

Task list of Semmelweis University in the project

1. Analysis of hits (druglikeness, synthetic feasibility, novelty) - *successfully completed*
2. Hits validated, validated hit list provided - *successfully completed*
3. Docking of small molecules - *successfully completed*
4. Lead candidate optimized and developed as well as selected for in vivo studies – *basic concept has been devalidated see final report to EC*
5. Lead series with proven efficacy in animal model – *4 .*

2019-01-01 - 2019-12-31

In accordance with the research plan of the consortium, our research group have collected the relevant scientific publications about alpha1-antitrypsin deficiency. We have received the preliminary screening data of IRE1 modulator molecules, and built a graphical database for computational screening. We used molecular modelling to explore the possible modes of action of the IRE1 activator compounds. The crystallographic structures of IRE1 were obtained from Protein Data Bank. IRE1 possesses both kinase and ribonuclease (RNase) activities. ADP, if bound to the nucleotide binding site, can activate IRE1's RNase. According to Wiseman et al. (10.1016/j.molcel.2010.04.001) the enzyme has at least two distinct ligand sites for activator compounds.
The nucleotide binding site is where natural activating ligand ADP exerts its effects. There are a number of known kinase inhibitors that can activate RNase in this manner. Since kinase active sites are highly conserved regions, it is reasonable to presume that compounds with kinase hinge region-binding motifs are potentially activators of IRE1 RNase as well. The crystal structures 3LJ1 and 3LJ2 are examples of IRE1 in complex with RNase activating kinase inhibitors at the nucleotide binding site. Both structures were used in our glide docking experiments.
The second binding site is an allosteric site at the dimer interface of IRE1's kinase extension nuclease (KEN) domain. Quercetin was found to activate IRE1 RNase, while potentiating ADP's activating effect as well. The crystallographic structure 3LJ0 revealed the unanticipated binding site, where the dimerization is strengthened by two ligand molecules facing each other, thereby forming π-π stacking interactions between them, and also the aromatic side chains of surrounding amino acids. Hydroxyl groups of quercetin interact with Glu988 and Ser984 via H-bonds. Related flavonol compounds also show RNase activating effects, although, to a lesser extent. 3LJ0 was also used in our glide docking experiments.
After examining our docking models, we expect at least 7 of the docked activators to bind to the nucleotide site. All compounds with a 2-(N-Anilino)pyrimidine core structure were found to form the anticipated H-bond interactions with both the peptide oxygen and the nitrogen of the hinge region residue Cys748. Docking scores correlate with this presumption, as the lowest scoring compounds (i.e. those that bind strongest according to the model) were those with an anilinopyrimidine structure or its close derivatives. These compounds are:
LDC206585
LDC207112
LDC001526
LDC041407
LDC000928
LDC191605
LDC001617
Docking into the quercetin active site was conducted in 2 steps. Quercetin from one chain was removed and ligands were docked in its position. Next, the remaining quercetin molecule was removed from the output structures, and each ligand was docked into their corresponding IRE1-ligand complex. We identified 4 ligands that are possibly capable of binding to the quercetin site, by formation of an H-bond with Glu988 and π-π stacking interactions with Phe1112 and the aromatic rings of the adjacent ligand molecule:
LDC206805
LDC206964
LDC162756
LDC206585 (probably a nucleotide site inhibitor)
Other compounds either lack a suitable H-bond donor, or do not fit well in the binding site. It is premature to state, that these molecules could not bind to the quercetin site, but they are less likely to provide the forces required to maintain the dimeric form of the protein.

SiteMap was used to identify a possible binding site other than the nucleotide and quercetin sites. A cavity was found between the two monomers in 3LJ0, although it is unclear if ligands would modify RNase activity if bound at this site. Docking indicates that hesperadin (LDC001741) has a suitable shape to fit into the cavity, although it is a known kinase inhibitor, and is more likely to occupy the nucleotide site when bound.

Further investigation is necessary to establish a binding site for the other compounds. Some of them are unusual structures in terms of drug-likeness, others lack any characteristic features to pinpoint their exact mode of action. Competition with ADP would prove the molecule to be a nucleotide site activator, while synergistic effect with ADP could indicate a binding site other than the nucleotide site; this could be examined with an ADP competition assay.
For the compounds we suspect to be nucleotide site inhibitors, the binding models can be used to aid the planning of new active analogues. The models are also suitable for searching for ligands, possibly IRE1 RNase activators. The docking of a library of ~25000 molecules is planned, and initial hits should be investigated by high precision docking, and eventually, tested in biological assays.

2020-01-01 - 2020-12-31

Molecular modelling was applied to a set of diverse molecules provided to us by LDC. RNase activation and XBP1 splicing values for these compounds have previously been measured in biological assays, with varying results. Docking into rigid protein structures of IRE1 (PDB: 3LJ0, 3LJ1, 3LJ2) provided some additional insight into the relation between binding to the IRE1 active site, the resulting effect on the RNase function, and on XBP1 splicing.
36 compounds were found to form at least 1 H-bond with the hinge region of the IRE1 kinase. Of these compounds, two (UDO031303, LDC084266) had a lower RNase AC50 value lower than 1 μM. The docking model for UDO031303 seems plausible, as the size of the molecule is optimal for entering and staying inside the ATP-binding cavity. It also forms two H-bonds (1 donor, 1 acceptor) with the hinge amino acid Cys748. The model and the compound itself may serve as a template for designing new compounds, or searching for similar existing compounds exhibiting greater effect on RNase activation and XBP1 splicing.
The allosteric Quercetin binding site (modelled in PDB 3LJ0) requires a pair of molecules bound at the same time, engaging in a stacking interaction with each other and nearby aromatic amino acid side chains. Therefore, simple docking experiments were insufficient in providing appropriate scoring of the docked compounds. The experiments indicate that ligands of the allosteric site should contain multiple (at least two) aromatic ring systems, connected by short (1-3 atoms length) chains or direct bonding, and at least one aromatic H-donor substituent.
To determine suitable scaffolds and identify possible IRE1 nucleotide binding site ligands, high-throughput virtual screening was applied to a molecular library of ~25000 compounds on multiple IRE1 crystallographic structures (PDB 3LJ1, 3LJ2, 4YZC, 6HX1, 6URC). A smaller group of molecules that scored lower than -7 kcal/mol were analysed further, by conducting various dockings with SP and XP precision. 35 hit compounds were chosen after consideration of docking scores, molecular weight, predicted or known pharmacological activity, and overall drug-likeness. The selected compounds were available from chemical provider. Solubility and lipophilicity of these compounds were determined by HPLC experiments using a method described by Z. Nemes et al. (Bioorg Med Chem Lett. 2018 Aug 1;28(14):2391-2398.). LogS at pH = 2 and pH = 7.4 were determined by comparing the AUC of DMSO solutions of known concentration (120 μM) to AUCpH=2.0 and AUCpH=7.4. Apparent lipophilicity was determined by retention times at pH = 12, at which most basic amine groups exist in deprotonated state. Retention times of compounds that are detected by HPLC and whose LogP are established in literature, were chosen as reference in the range of LogP = -0.07 to 3.26. Solubility and lipophilicity data are summarized in Table 1.
QikProp was used to predict logIC50 for hERG K+-channel blockage, Caco-2 and MDCK permeability in nm/s and percentage of human gastrointestinal absorption. Data from the in silico

measurements for ligands at predicted protonation state in pH = 7.4 environment are shown in Table 2. The values were also calculated for pH = 2, pH = 12 and for unionized form of the ligands. Based on the obtained acceptable ADME parameters of the compounds, their further testing in biological assays at LDC (Lead Discovery Center) looks like promising.

Table 1: Solubility and lipophilicity data for the 35 test compounds

#	LogS2.0	LogS7.4	LogD12	IUPAC name
1	-3,927	-3,930	1,63	1-(2-deoxypentofuranosyl)-5-pentylpyrimidine-2,4(1H,3H)-dione
2	-4,088	-4,442	2,26	4-fluoro-2-methoxy-5-{[4-(2-methoxyphenyl)pyrimidin-2-yl]amino}benzamide
3	-4,752	-3,534	2,00	N-({(3Z)-3-[(4-{[(1S)-1-benzyl-2-methoxy-2-oxoethyl]carbamoyl}-3,5-dimethyl-1H-pyrrol-2-yl)methylene]-2-oxo-2,3-dihydro-1H-indol-5-yl}carbonyl)-D-phenylalanine
4	-5,342	-5,312	2,66	(2R)-2-(3-chlorophenyl)-N-[(3Z)-3-{[3,5-dimethyl-4-(3-oxo-3-piperidin-1-ylpropyl)-1H-pyrrol-2-yl]methylene}-2-oxo-2,3-dihydro-1H-indol-5-yl]-2-hydroxyacetamide
5	-4,011	-4,069	1,57	N-{3-[(4-{6-[(2-aminoethyl)amino]pyridin-3-yl}pyrimidin-2-yl)amino]phenyl}-2-morpholin-4-ylacetamide
6	-4,105	-4,843	2,17	(2R)-2-(3-chlorophenyl)-N-[(3Z)-3-({3,5-dimethyl-4-[3-(4-methylpiperazin-1-yl)-3-oxopropyl]-1H-pyrrol-2-yl}methylene)-2-oxo-2,3-dihydro-1H-indol-5-yl]-2-hydroxyacetamide
7	-3,944	-3,941	1,96	2-allyl-1-[6-(1-hydroxy-1-methylethyl)pyridin-2-yl]-6-{[4-(4-methylpiperazin-1-yl)phenyl]amino}-1,2-dihydro-3H-pyrazolo[3,4-d]pyrimidin-3-one
8	-4,001	-4,390	1,90	1-(3-{8-[(3,4-dimethoxyphenyl)amino]imidazo[1,2-a]pyrazin-3-yl}phenyl)-3-(2-hydroxyethyl)urea
9	-3,653	-3,863	1,01	3-(5-{(Z)-[5-(L-lysylamino)-2-oxo-1,2-dihydro-3H-indol-3-ylidene]methyl}-2,4-dimethyl-1H-pyrrol-3-yl)propanoic acid
10	-4,814	-3,918	1,65	(2R)-2-[({(3Z)-3-[(4-{[(1S)-4-methoxy-1-(methoxycarbonyl)-4-oxobutyl]carbamoyl}-3,5-dimethyl-1H-pyrrol-2-yl)methylene]-2-oxo-2,3-dihydro-1H-indol-5-yl}carbonyl)amino]-3-phenylpropanoic acid
11	-3,964	-3,976	1,49	4-{[5-amino-1-(2,6-difluorobenzoyl)-1H-1,2,4-triazol-3-yl]amino}benzenesulfonamide
12	-4,034	-4,055	1,78	(3Z)-3-(3,4-dihydroxybenzylidene)-N-[(1R)-1-(4-methoxyphenyl)ethyl]-2-oxoindoline-5-sulfonamide
13	-3,956	-4,704	1,14	(3Z)-N-[(1R)-1-(3-methoxyphenyl)ethyl]-2-oxo-3-(3,4,5-trihydroxybenzylidene)indoline-5-sulfonamide
14	-5,618	-5,068	2,59	Dimethyl N-({5-[(Z)-(5-{[(1R)-1-benzyl-2-methoxy-2-oxoethyl]carbamoyl}-2-oxo-1,2-

				dihydro-3H-indol-3-ylidene)methyl]-2,4-dimethyl-1H-pyrrol-3-yl}carbonyl)-D-glutamate
15	-3,609	-3,676	1,42	2-{[4-(2-aminoethoxy)phenyl]amino}-8-methylpyrido[2,3-d]pyrimidin-7(8H)-one
16	-3,943	-4,062	2,27	2-{[4-(2-aminoethoxy)phenyl]amino}-8-bicyclo[2.2.1]hept-2-ylpyrido[2,3-d]pyrimidin-7(8H)-one
17	-3,802	-3,871	1,59	2-{[4-(2-aminoethoxy)phenyl]amino}-8-ethylpyrido[2,3-d]pyrimidin-7(8H)-one
18	-4,059	0	2,68	1-(3-chlorophenyl)-3-{3-[(5,6-dimethyl-7H-pyrrolo[2,3-d]pyrimidin-4-yl)amino]phenyl}urea
19	-3,919	-4,317	1,66	9-(3,5-dihydroxybenzoyl)-9H-carbazole-2,3,6,7-tetrol
20	-3,993	-6,207	2,52	3-(3-aminophenyl)-N-1,3-benzodioxol-5-ylimidazo[1,2-a]pyrazin-8-amine
21	-3,928	-4,022	0,94	N-[(3Z)-3-(1H-imidazol-5-ylmethylene)-2-oxo-2,3-dihydro-1H-indol-5-yl]-L-alaninamide
22	-4,017	-4,163	2,38	(3Z)-3-({3,5-dimethyl-4-[(1-methylpiperidin-4-yl)carbamoyl]-1H-pyrrol-2-yl}methylene)-N-[1-(4-fluorophenyl)propyl]-2-oxoindoline-5-carboxamide
23	-3,911	-3,957	2,02	N-[5-(2-{[3-(trifluoromethyl)phenyl]amino}pyrimidin-4-yl)pyridin-2-yl]ethane-1,2-diamine
24	-4,036	-3,989	2,14	N-(5-{2-[(2-fluorophenyl)amino]pyrimidin-4-yl}pyridin-2-yl)ethane-1,2-diamine
25	-3,935	-4,043	1,73	N-(5-{2-[(2-chlorophenyl)amino]pyrimidin-4-yl}pyridin-2-yl)ethane-1,2-diamine
26	-3,987	-4,287	2,33	(3Z)-3-({3,5-dimethyl-4-[(2-pyrrolidin-1-ylethyl)carbamoyl]-1H-pyrrol-2-yl}methylene)-N-(3-methoxybenzyl)-2-oxoindoline-5-carboxamide
27	-4,866	-4,887	2,42	N-(5-{2-[(3-chlorophenyl)amino]pyrimidin-4-yl}pyridin-2-yl)ethane-1,2-diamine
28	-3,912	-4,335	2,53	N-(5-{2-[(4-chlorophenyl)amino]pyrimidin-4-yl}pyridin-2-yl)ethane-1,2-diamine
29	-3,936	-4,558	2,13	N-[5-(2-{[4-(4-methylpiperazin-1-yl)phenyl]amino}pyrimidin-4-yl)pyridin-2-yl]ethane-1,2-diamine
30	-5,100	-3,915	1,34	N-{5-[2-({4-[2-(dimethylamino)ethoxy]phenyl}amino)pyrimidin-4-yl]pyridin-2-yl}ethane-1,2-diamine
31	-3,904	-4,790	1,95	8-cyclopentyl-2-({4-[2-(methylamino)ethoxy]phenyl}amino)pyrido[2,3-d]pyrimidin-7(8H)-one
32	-3,606	-3,667	1,91	N-({5-[3-(4,6-difluoro-1H-benzimidazol-2-yl)-1H-indazol-5-yl]-4-methylpyridin-3-yl}methyl)ethanamine
33	-3,947	0	2,90	N-[(1S)-2-amino-1-(2-chlorobenzyl)ethyl]-1-[2-(methylamino)pyrimidin-4-yl]-1H-pyrazole-4-carboxamide
34	-3,901	-4,815	1,48	2-amino-1-(4-chlorophenyl)-1-[5-(1H-pyrrolo[2,3-b]pyridin-4-yl)-2-thienyl]ethanol

35	-4,008	-5,418	3,59	(1S)-2-amino-1-(5-chloro-2-thienyl)-1-[4-(1H-pyrrolo[2,3-b]pyridin-4-yl)phenyl]ethanol

Table 2: In silico measurements for hERG blockage, Caco-2 and MDCK cell permeability and percentage of human gastrointestinal absorption of the test compounds

#	logIC50	hERG	Caco-2	MDCK % absorption
1	-3.718	136.971	57.696	67.227
2	-6.325	528.991	423.469	93.291
3	-6.036	6.389	2.671	40.256
4	-5.880	94.318	154.856	74.132
5	-8.770	16.120	6.990	56.423
6	-6.241	21.063	36.018	54.897
7	-7.511	220.534	106.808	76.572
8	-5.508	203.018	127.156	83.976
9	-5.804	0.094	0.034	0.000
10	-5.814	1.661	0.623	21.979
11	-6.156	28.455	27.698	56.604
12	-5.091	53.064	21.103	67.796
13	-6.404	16.107	5.770	57.522
14	-7.996	32.679	12.259	48.885
15	-6.484	78.531	34.986	67.395
16	-7.043	212.225	102.465	85.094
17	-6.349	94.939	42.951	70.675
18	-5.633	514.097	826.029	100.000
19	-5.306	5.629	1.832	26.406
20	-6.081	1037.835	514.971	100.000
21	-6.108	18.297	7.245	48.525
22	-7.972	100.376	82.332	69.110
23	-6.003	1150.374	575.585	100.000
24	-9.178	36.445	16.881	79.085
25	-7.543	10.298	4.306	44.735
26	-7.861	112.400	51.549	78.249
27	-6.022	98.875	66.548	71.633
28	-6.923	115.530	53.103	77.283
29	-4.405	407.158	293.946	85.840
30	-4.528	0.836	0.377	10.496
31	-5.538	249.035	145.076	87.687
32	-7.051	258.733	126.938	89.015
33	-6.592	915.553	449.712	100.000
34	-5.284	902.770	442.929	100.000
35	-8.709	224.029	267.755	85.489

Questions of the reviewer about work of period 2020-01-01 - 2020-12-31:

In this interim report, the PI provided in silico data of 35 compounds in two Tables (Table 1: Solubility and lipophilicity; Table II: ADME properties - hERG K+-channel blockage, Caco-2 and MDCK permeability and human gastrointestinal absorption at different ionization form of ligands), however, there are no information given about the number and type of derivatives with "acceptable" ADME parameters. Principal investigator is invited to provide further details of biological evaluation,

i.e. the list of assays and the meaning of "looks like promising". The approach of applying Lead Discovery Center to validate in silico results is questionable. There are no publication attached to the annual report.

Answers to the reviewer's questions:

This is an interim report, the project's end date just have been prolonged by 180 days from the original end date (to 29.05.2022) due to the Covid pandemic.
Table 1 contains experimental (HPLC) logS and logP data, determined according to Z. Nemes et al. Bioorg Med Chem Lett. 2018, 28(14):2391-2398. Range of recommended values for logP: -2-6.5, for logS: -6.5-0.5 . (S in mol/dm^3). The compounds are all fitting into the druglike range.
Table 2 contains in silico determined ADMET properties. These calculations assumed passive transport only. Acceptable ADMET parameters are the following: In case of hERG blockage values below -5 can be problematic, here in most of cases these are predicted as an average value about -6, which looks like poor, but hERG inhibition is much more difficult than to characterize it with one numeric value. Caco-2 cell permeability in nm/sec for non-active transport: <25 poor, >500 great. Nine compounds in Table 2 were predicted poor, the others were acceptable or great. MDCK cell permeability in nm/sec are considered to be a good mimic for the blood brain barrier, for non-active transport: <25 poor, >500 great. 11 compounds were predicted poor, the others were acceptable. The Human Oral Absorption value >80% high, <25% poor. 32 compounds were predicted having acceptable HOA, and 3 compounds were predicted poor. The examined compounds had excellent scores in the IRE1 in silico docking, showed druglikeness (even not in every parameters), and these compounds were available for experimental studies.
At the LDC the following ADME Assays are running routinely: Metabolic stability in Liver microsomes: Phase I Metabolism: CLint Phase I HLM (human), MLM (mouse), RLM (rat) GLM (guinea pig), Phase II metabolism: %Remain Phase II HLM, MLM, RLM Plasma Stability: PlasmaStab human, mouse, rat, Plasma Protein binding: PPB human, mouse, rat, Caco-2 Permeability assays: Caco-2 (bidirectional), CYP Inhibition: CYP1A2, CYP2C9, CYP2C19, CYP2D6, CYP3A4, Chemical Stability, Solubility (equilibrium, from powder), GSH adduct formation, DEREK in silico tox prediction., Cytotox PBMC assay, MitoTox GluGal, HepatoTox Assays in HepG2 Liver cells (mouse, rat, human),CardioTox hERG binding assay, GenoTox GreenScreen Assay. The LDC can test any of the compounds in the above mentioned assays; however most important will be testing of the compounds for activation of IRE1, which we can do in cell based assays of XBP-1-splicing, or direct IRE1 activation assay. On top of this Eric Chevet (consortium member of INSERM) has IRE1 activation assays up and running, he would be also able to test the hit list from the virtual screening. So far, we have not published anything from the ERAAT results, and we are planning not to publish anything in the near future, as this would undermine our position to file patents on the IRE1 Activators, which will be essential for out licensing and further development of the IRE1 Activators.

2021-01-01 - 2021-12-31

Previously, 35 virtual hits were selected for testing on IRE1 according to docking experiments, experimental solubility and lipophilicity data and ADME property prediction. Compounds that had impurities exceeding 5% were purified by HPLC. The compounds were provided to LDC for assessment of their effect on human IRE1 in a cell free RNase activity assay, and were tested in a

dilution series ranging from 2.3 nM to 5 µM in 3-fold dilution steps. Two compounds have exerted remarkable effect on IRE1 RNase activity, one of which was identified as an inhibitor (#9), while the other one as an activator (#16). Another compound was determined as a weak activator (#25), and several others have also had some effect on RNase activation, especially at 5 µM of concentration, although the difference of the RNase activity relative to the control did not exceed 20%. In the case of the inhibitor compound 9, the RNase activity relative to the control was measured at 46.0% at 5 µM, and 78.7% at 1.67 µM. For compound 16, the RNase activity was measured at 344.9% at 5 µM and 125.3% at 1.67 µM. Three compounds (#9, #16, #25) were identified as hits, two of which have a potent effect on RNase activity, out of the 35 molecules selected by virtual screening. This newly obtained biological data underlines the success of our in silico experiments.
Both compound 9 and 16 have a molecular structure similar to known kinase inhibitors, and are capable of binding to kinase hinge regions, and thus, were identified as ligands of the ATP binding pocket. Their effect on IRE1 were further analyzed through High precision docking experiments (Induced Fit docking - Pose Prediction mode) and Molecular Dynamics simulations of the output structures. Newer IRE1 crystallographic structures (PDB IDs: 6W3K, 6XDB, 6XDD) were also used in Glide docking experiments. Three IRE1 structures (yeast, 3LJ1; human, 6W3K and 6XDD) were chosen for the Induced Fit dockings by considering ligand poses and respective docking scores. Induced Fit results were evaluated by assessment of the crucial interactions required for binding to the ATP pocket. The models of the hit compounds in complex with the 6XDD IRE1 structure provided the most appealing results, and therefore, the output structures were chosen as the starting point for the Molecular Dynamics simulations. 20 ns simulations were carried out using SPC water model and 0.15M NaCl salt concentration. Persistent interactions of Compound 9 with the binding site are H-bonds with Glu643 and Cys645 of the hinge region (Glu746 and Cys748 in yeast IRE1), both direct H-bond and water bridge connection with Asp711 of the DFG-motif, ionic bond with the catalytic Lys599 residue through a chloride ion, and H-bond with Arg588. Interactions of Compound 16 are two H-bonds formed with Cys645 (donor and acceptor), water bridge formed with Asp711, and H-bond with Glu651. The most pronounced difference between the simulations is the effect on the αC Helix of the kinase, as the inhibitor Compound 9 seems to disrupt the helical structure of the αC, while Compound 16 rather seems to stabilize the active kinase conformation. An indicator of the active conformation is the presence of the Lys599-Glu612 interaction, (https://doi.org/10.1073/pnas.1906999116, https://doi.org/10.18632/oncotarget.3864) and the disruption of this interaction has been cited as a mechanism of action for inhibitors previously published (https://doi.org/10.1124/mol.115.100917). In the MD simulation of Compound 9 a large (10-15 Å) gap is formed between these residues due to the conformational change in the αC helix, while direct interaction is observed in simulations of Compound 16. It has been implicated that most ATP-competitive inhibitors interacting with the kinase ATP pocket are activators of the RNase, and inhibition of the RNase observed for other compounds is due to direct or indirect disruption of the αC helix structure, hindering IRE1 dimer formation and preventing autophosphorylation and subsequent activation (https://doi.org/10.1021/acschembio.5b00940). These findings about the newly tested compounds are in accordance with literature on IRE1 and provide important knowledge about the nature of RNase activity modulators. The experiments will be of aid for our partners in selecting compounds for modulation of IRE1 RNase activity specifically as inhibitors or activators, as well as providing a scheme for synthesizing new analogues based on these scaffolds.

Questions of the reviewer about the work period 2021-01-01 - 2021-12-31

The research group identified 35 virtual hits by docking a database to the active site of inositol requiring enzyme 1 (IRE1) and post-processing primary hits by physchem and ADME filters. These compounds were tested against human IRE1 in a cell free RNase activity assay. Three compounds showed activity out of which compound 9 inhibited, while compounds 16 and 25 activated the RNase activity. The calculated hit rate of 8.5% suggests the developed virtual screening protocol successful. Next investigations were focused to the confirmed hits. Since compound 25 was a weak activator,

further studies included compounds 9 and 16 only. Their binding modes were first analyzed by induced fit docking and the best models were subjected to MD simulations. These calculations revealed their molecular mechanism of action. An obvious limitation of these compounds is their weak activity that should be optimized before evaluating them in more sophisticated disease biology models. The research group should focus to this activity in the next year. Nevertheless, hit finding would be worth to publish.

Answers to the questions of the reviewer:

The publication of our results requires the permission of our consortial partners. Especially, McGill and INSERM has patents in the topic of the project already. Use of any information from the patents requires their permission (even these are public, in accordance with the EC rules). After finishing the project the consortium will decide on the patenting option. Preparation of a patent will take about half- to one year. Any publication before the patent application would ruin the novelty of the patent. In case the consortium will not vote for patenting, we plan to publish the in silico molecular modeling and potential hit selection in a chemical informatics journal, e.g. in ACS Journal of Chemical Information and Modeling, as soon as possible.

Prof. Dr. László Őrfi
Semmelweis University