Development of small molecule modulators of stem cell reprogramming

Dissertation
For the achievement of the academic degree of the
Doctor in Natural Sciences
(Dr. rer. nat.)

Submitted to
Faculty of Chemistry and Chemical Biology
Technical University Dortmund
by

Andrei Ursu

From Bacau, Romania

Dortmund 2015

Development of small molecule modulators of stem cell reprogramming

Dissertation
zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

eingereicht an
der Fakultät für Chemie und Chemische Biologie der
Technischen Universität Dortmund
vorgelegt von

Andrei Ursu
Aus Bacau, Rumänien

Dortmund 2015

Eidesstattliche Versicherung (Affidavit)

Ursu Andrei
Name, Vorname
(Surname, first name)

154486
Matrikel-Nr.
(Enrolment number)

Belehrung:

Wer vorsätzlich gegen eine die Täuschung über Prüfungsleistungen betreffende Regelung einer Hochschulprüfungsordnung verstößt, handelt ordnungswidrig. Die Ordnungswidrigkeit kann mit einer Geldbuße von bis zu 50.000,00 € geahndet werden. Zuständige Verwaltungsbehörde für die Verfolgung und Ahndung von Ordnungswidrigkeiten ist der Kanzler/die Kanzlerin der Technischen Universität Dortmund. Im Falle eines mehrfachen oder sonstigen schwerwiegenden Täuschungsversuches kann der Prüfling zudem exmatrikuliert werden, § 63 Abs. 5 Hochschulgesetz NRW.

Die Abgabe einer falschen Versicherung an Eides statt ist strafbar.

Wer vorsätzlich eine falsche Versicherung an Eides statt abgibt, kann mit einer Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft werden, § 156 StGB. Die fahrlässige Abgabe einer falschen Versicherung an Eides statt kann mit einer Freiheitsstrafe bis zu einem Jahr oder Geldstrafe bestraft werden, § 161 StGB.

Die oben stehende Belehrung habe ich zur Kenntnis genommen:

Official notification:

Any person who intentionally breaches any regulation of university examination regulations relating to deception in examination performance is acting improperly. This offence can be punished with a fine of up to EUR 50,000.00. The competent administrative authority for the pursuit and prosecution of offences of this type is the chancellor of the TU Dortmund University. In the case of multiple or other serious attempts at deception, the candidate can also be unenrolled, Section 63, paragraph 5 of the Universities Act of North Rhine-Westphalia.

The submission of a false affidavit is punishable.

Any person who intentionally submits a false affidavit can be punished with a prison sentence of up to three years or a fine, Section 156 of the Criminal Code. The negligent submission of a false affidavit can be punished with a prison sentence of up to one year or a fine, Section 161 of the Criminal Code.

I have taken note of the above official notification.

Dortmund, 20.08.2015
Ort, Datum
(Place, date)

Unterschrift
(Signature)

Titel der Dissertation:
(Title of the thesis):

Development of small molecule modulators of stem cell reprogramming

Ich versichere hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Titel selbstständig und ohne unzulässige fremde Hilfe angefertigt habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie wörtliche und sinngemäße Zitate kenntlich gemacht.
Die Arbeit hat in gegenwärtiger oder in einer anderen Fassung weder der TU Dortmund noch einer anderen Hochschule im Zusammenhang mit einer staatlichen oder akademischen Prüfung vorgelegen.

I hereby swear that I have completed the present dissertation independently and without inadmissible external support. I have not used any sources or tools other than those indicated and have identified literal and analogous quotations.

The thesis in its current version or another version has not been presented to the TU Dortmund University or another university in connection with a state or academic examination.*

*Please be aware that solely the German version of the affidavit ("Eidesstattliche Versicherung") for the PhD thesis is the official and legally binding version.

Dortmund, 20.08.2015
Ort, Datum
(Place, date)

Unterschrift
(Signature)

1st Examiner
Prof. Dr. Herbert Waldmann
Department of Chemistry and Chemical Biology
Technical University Dortmund, Dortmund, Germany
and
Department of Chemical Biology
Max Planck Institute of Molecular Physiology, Dortmund, Germany

2nd Examiner
Prof. Dr. Hans R. Schöler
Medical Faculty
University of Münster, Münster, Germany
and
Department of Cell and Developmental Biology
Max Planck Institute for Molecular Biomedicine, Münster, Germany

Acknowledgements

I would like to thank Professor Dr. Herbert Waldmann for accepting me in his group and for allowing me to work on this very interesting and challenging project. I am grateful for his continuous support, supervision and helpful suggestions throughout the entire doctoral study. Moreover, I appreciate the degree of research freedom I enjoyed during my time spent at the Max Planck Institute of Molecular Physiology, Dortmund. Under your guidance, I transformed myself from a chemist into a chemical-biologist and beyond.

I owe my sincere thanks to Dr. Slava Ziegler, for always having her office door open for suggestions, very fruitful scientific discussions and for encouraging me to explore the biological world.

I am grateful to my beloved fiancée and future wife, Alina Homorozean, for the continuous moral support, understanding and accepting that working over the weekend is an integral part of science.

I acknowledge my former and present colleagues from the Technical University Dortmund, especially Dr. Jakub Švenda, Dr. Sara-Lopez Tosco, Chantale Martin, Dr. Rajarshi Samanta and Dr. Rishikesh Narayan for the long scientific and non-scientific discussions. I am grateful to the entire chemistry department of the Technical University Dortmund for the nice atmosphere and working environment.

This work would not have been possible without the excellent research atmosphere at the Max Planck Institute of Molecular Physiology, Dortmund. Here I would like to mention my former and current colleagues Dr. Christian Spiteri, Dr. Yao Wang and Javier Ceballos for insightful suggestions related to the organic synthesis part of this study. I have been fortunate in getting the chance to work with dedicated biologists and chemical-biologists. Here I would like to mention Dr. Shobhna Kapoor, Dr. Yasemin Akbulut, Dr. Gloria-Vendrell Navarro, Sumersing Patil and Lea Kremer. I am thankful to Professor Dr. Hans R. Schöler, Dr. Damir J. Illich and Dr. Miao Zhang from the Max Planck Institute for Molecular Biomedicine (Münster, Germany) for the productive collaboration. I want to express my gratitude to Dr. Yasushi Takemoto (RIKEN-Max Planck Joint Research Center for Systems Chemical Biology, Wako, Japan) and Arthur Porfetye (Max Planck Institute of Molecular Physiology, Dortmund, Germany) for supporting me with the optimization of the kinase assay and performing the crystallization studies, respectively.

I acknowledge the International Max Planck School for Chemical and Molecular Biology (IMPRS-CMB) for granting financial support during the doctoral study. In particular, I am thankful to Dipl. Biol. Christa Hornemann and Prof. Dr. Andrea Musacchio for the very good organization and coordination of the graduate school.

I am grateful to Professor Dr. Bert Vogelstein (Sidney Kimmel Comprehensive Cancer Center, Baltimore, USA) and Professor Dr. Kenneth W. Kinzler (Ludwig Center at Johns Hopkins University, Baltimore, USA) for proving the wild type and the *TP53* mutant HCT116 colorectal carcinoma cell lines. I acknowledge the Dortmund Protein Facility (DPF) for the plasmid construction, expression and purification of the CSNK1D/E isoenzymes and the High Resolution Mass Spectrometry (HRMS) team for the MS analyses.

Table of Contents

Acknowledgements

Abstract I

Zusammenfassung III

Basic definitions 1

1. Introduction 2
  - 1.1. Stem cell research 2
    - 1.1.1. Embryonic development 2
    - 1.1.2. (Pluripotent) stem cell states 3
    - 1.1.3. Integrative genetic methods to access pluripotent cellular states 6
    - 1.1.4. Non-integrative genetic methods to access pluripotent cellular states 8
    - 1.1.5. Small molecules in stem cell research 9
    - 1.1.6. Chemical reprogramming 12
    - 1.1.7. Potential applications of stem cell technologies 14
    - 1.1.8. Differences between mESCs and mEpiSCs 16
    - 1.1.9. The mEpiSC population is heterogeneous 18
    - 1.1.10. Reprogramming mEpiSCs to mES-like cells – state of the art 19
    - 1.1.11. Reprogramming late-stage mEpiSCs to mES-like cells – preliminary results 21
  - 1.2. Chemical Genetics 22
    - 1.2.1. Definition 22
    - 1.2.2. Chemical versus genetic intervention to probe biology 24
    - 1.2.3. Forward chemical genetics applied to stem cell research 27
    - 1.2.4. Chemical genetics and stem cells – future challenges 28

2. Aim of the Thesis 30

3. Experimental part 32
  - 3.1. Experimental part – chemistry 32
    - 3.1.1. Materials 32
    - 3.1.2. Methods 32
  - 3.2. Experimental part - Biology 39
    - 3.2.1. Materials 39
    - 3.2.1.1. Cell lines 40
    - 3.2.2. Methods 41
    - 3.2.2.1. Reprogramming assay using mouse epiblast stem cells 41
    - 3.2.2.2. Immunocytochemistry 41
    - 3.2.2.3. Preparation of cell lysate for detecting CSNK1A1/D/E 42
    - 3.2.2.4. Immunoblotting 42
    - 3.2.2.5. Plasmid construction, protein expression and purification 43

3.2.2.6. Crystal structure determinations 44
3.2.2.7. Cell lysate preparation for target engagement and chemical proteomics 46
3.2.2.8. Kinase profiling at Eurofins (former Milipore) 46
3.2.2.9. Determination of CSNK1D/E activity using the KinaseGlo® Plus luminescent assay 47
3.2.2.10. CSNK1A profiling at Reaction Biology 48
3.2.2.11. Detection of target engagement using the ActivX™ ATP probe 48
3.2.2.12. Chemical proteomics experiment based on the ActivX™ ATP probe 49
3.2.2.13. Mass spectrometric analysis 50
3.2.2.14. WST-1 cell proliferation assays 51
3.2.2.15. CellTiter-Glo® viability assay 52
3.2.2.16. CellTox™ Green cytotoxicity assay 53
3.2.2.17. Tryphan blue exclusion test of cell viability 53
4. Results 54
4.1. Structure activity relationship (SAR) studies 54
4.2. Characterization of the reprogramming activity 60
4.3. Pteridine-based analogs inhibit the CSNK1 kinase family 63
4.4. Screening of TR-based derivatives yields potent and / or selective CSNK1D/E inhibitors 70
4.5. Evaluation of the selected pteridine analogues against CSNK1A1 77
4.6. Crystallization studies 80
4.7. Detection of target engagement for Epiblastin A 85
4.8. Revealing novel interaction partners of Epiblastin A by means of chemical proteomics 90
4.9. Data analysis of chemical proteomics experiments 94
4.10. Kinase profiling reveals targets of Epiblastin A potentially involved in reprogramming 95
4.11. Investigation of Epiblastin A in cancer cell lines 98
4.11.1. Epiblastin A inhibits the proliferation of wt HCT116 colorectal carcinoma cells 98
4.11.2. Growth inhibition of HCT116 cells by Epiblastin A is influenced by various parameters 104
4.11.3. Epiblastin A affects the growth of various cancerous and non-cancerous cell lines 106
5. Discussion 108
6. Summary 118
7. Outlook 122
8. Appendix 124
9. Abbreviations 125
10. References 127
11. Compound characterization 137
Curriculum Vitae

Abstract

Accessing a large number of high quality pluripotent stem cells from differentiated cells by means of small molecule treatment, i.e. chemical reprogramming, is a prerequisite to bring stem cell technologies a step closer to clinical applications. Triamterene (**TR**), a pteridine-based analog, has been previously shown to convert late-stage mouse epiblast stem cells (mEpiSCs) into embryonic stem (ES)-like cells. As a continuation of the efforts to further improve the reprogramming activity, a **TR**-based compound collection was synthesized that culminated in the identification of Epiblastin A, which is approximately 2.3 fold more active compared to the initial hit. Since genetic and chemical inhibition of the casein kinase 1 (CSNK1) family has been previously identified as the driving force of the reprogramming process, the inhibitory activity of the compound collection against CSNK1 isoenzymes was investigated using a screening platform based on KinaseGlo® reagent. This study revealed structural features required for potent and / or selective interaction of the **TR**-analogs and these isoforms. Furthermore, target engagement of Epiblastin A with CSNK1A1/D/E isoenzymes was demonstrated in cell lysates obtained from HCT116 colorectal carcinoma cell line using the ActivX™ ATP probe. Moreover, additional targets addressed by Epiblastin A in HCT116 cell lysate were disclosed by means of chemical proteomics strategy using the same reagent. Among those, STK10, PIP4K2A and PIP4K2C kinases were identified and in addition STK10 was also shown to interact with Epiblastin A in cell lysates. Finally, an *in vitro* profiling study against 123 kinases furnished further kinase targets such as BRSK1, EEF2K, EGFR, MKNK2 and RIPK2, thus revealing potential modulators that might be involved in the chemical reprogramming mechanism.

Zusammenfassung

Die chemisch induzierte Reprogrammierung von pluripotenten Stammzellen ist ein wichtiger Schritt in der Entwicklung von klinisch relevanten Methoden und Anwendungen von Stammzellen. Die Verfügbarkeit von qualitativ hochwertigen pluripotenten Stammzellen bildet dafür derzeitig eine wichtige Schlüsselstelle. In der Vergangenheit es wurde gezeigt dass das Pteridin-basierte Triamteren (**TR**) in der Lage ist Stammzellen aus Maus Epiblasten (mEpiSCs) auch aus späten Entwicklungsstadien in embryonale Stammzellen (ES) zu konvertieren. Im Zuge der **TR**-Optimierung in Hinblick auf eine gesteigerte Reprogrammierungsaktivität wurde Epiblastin A aus einer **TR**-inspirierten Substanz-bibliothek identifiziert. Epiblastin A zeigt eine 2.3-fach erhöhte Aktivität im Vergleich zum Primärhit und stellt die Grundlage für die hier weiter aufgeführten Studien. Die Ergebnisse führten zur Hypothese, dass die Casein-Kinase (CSNK1) ein wichtiges potenzielles Schlüsselenzym für den von Epiblastin A induzierten Reprogrammierungsprozess ist. Hierzu wurden **TR**-Analoga in einem KinaseGlo® basiertem Testsystem auf CSNK1-modulierende Aktivität getestet. Daraus wurde eine Strukturaktivitätsanalyse (SAR) aufgestellt. Im Folgenden konnte in zellbasierten Experimenten bestätigt werden, dass Epiblastin A die Aktivität der CSNK1A1/D/E Isoenzyme inhibiert. Hierzu wurde im komplexen Zelllysat der Darmkrebszelllinie (HCT116) die Wechselwirkung von Epiblastin A mit CSNK1A1/D/E Isoformen untersucht. In zusätzlichen Studien, die auf einem klassisch chemisch-proteomischen Versuchsaufbau basierten, wurden aus HCT116-Lysat zusätzliche neue Zielproteine von Epiblastin A identifiziert. Darauf basierend wurde eine Interaktion mit den Kinasen STK10, PIP4K2A und PIP4K2C postuliert und in weiterführenden zelllysatbasierten Studien konnte STK10 als neuer Interaktionspartner von Epiblastin A bestätigt werden. In einer zusätzlichen, weitangelegten *in vitro* Studie wurden insgesamt 123 Kinasen als mögliche Zielproteine von Epiblastin A getestet. Hierbei wurden zusätzlich BRSK1, EEF2K, EGFR, MKNK2 und RIPK2 als weitere potenzielle Interaktionspartner im Epiblastin A induzierten Reprogrammierungsprozess von mEpiSC identifiziert.

Basic definitions

Cell differentiation = defines the process whereby a cell changes towards a more specialized cell type, for example a pluripotent cell changes into somatic cell.

(Stem) cell potency = defines the ability of a stem cell to differentiate into other cell types; the more cell types a stem cell can differentiate into, the greater its cellular potency.

Totipotent stem cells = these cells can give rise to an entire viable organism; they can differentiate into embryonic as well as extraembryonic tissues that are required for the embryo development.

Pluripotent stem cells = these cells can differentiate into cells derived from the three germ layers (ectoderm, mesoderm and endoderm).

Multipotent stem cells = these cells can solely give rise to limited cell types mostly from related lineages.

Oligopotent stem cells = stem cells that can differentiate into just few cell types.

Unipotent stem cells = stem cells that can differentiate into a single cell type, namely their own.

Cellular reprogramming / dedifferentiation = process upon which a specialized cell type transitions to a less specialized cellular state, i.e. a cell with limited differentiation potential is converted into a cellular state with increased differentiation potential.

Self-renewal = cellular process whereby the identity of a stem cell is indefinitely maintained upon multiple cell divisions.

Teratoma formation = experimental procedure assessing the ability of stem cells to form tumors containing differentiated cells derived from the three germ layers (ectoderm, mesoderm and endoderm).

Transdifferentiation / lineage reprogramming = transition process of a cell type into another cell type without undergoing a pluripotent stem-like state.

1. Introduction

1.1. Stem cell research

1.1.1. Embryonic development

Embryo development occurs in a stepwise and highly coordinated fashion. The process begins with the oocyte fertilization by the sperm that yields the zygote at 0.5 dpc (dpc = days post conception and suggests the age of an embryo; hereafter referred to as E0.5), a cell that is characterized by *totipotency* (see **Figure 1**). Subsequent cell divisions generate the early blastocyst at the pre-implantation stage (E3.5). At this point, embryonic stem cells (ESCs) can be isolated from the inner cell mass (ICM).

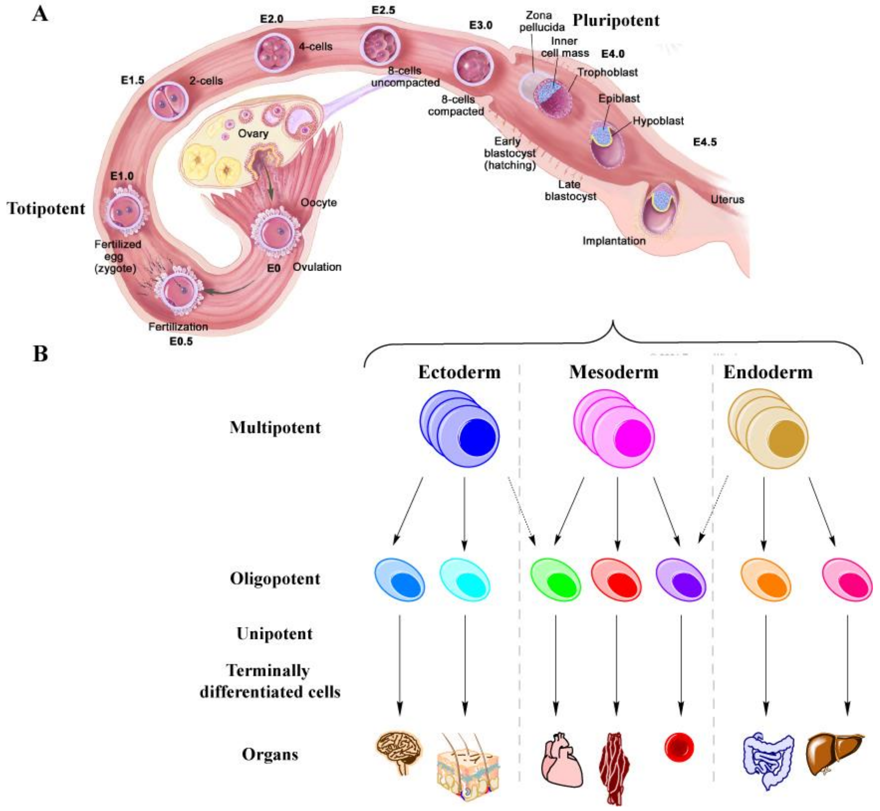

A
E2.0
4-cells
E2.5
8-cells uncompacted
E3.0
8-cells compacted
Zona pellucida
Pluripotent
Inner cell mass
E4.0
Trophoblast
Epiblast
Hypoblast
E4.5
E1.5
2-cells
Ovary
Early blastocyst (hatching)
Late blastocyst
Uterus
E1.0
Totipotent
Oocyte
Fertilized egg (zygote)
E0
Ovulation
Implantation
Fertilization
E0.5
B
Ectoderm
Mesoderm
Endoderm
Multipotent
Oligopotent
Unipotent
Terminally differentiated cells
Organs

Figure 1. Development of the mouse embryo from embryonic day 0 (E0) to organ formation. (A) Stages of the embryo development in mouse from embryonic day 0 (E0) to day 5 (E5) (© 2001 Terese Winslow). (B) Hierarchical stem cell specification process (adapted after **Figure 1** in Wobus et al. Physiol. Rev. **2005**, *85*, 635-678).

These cells are descendants of the zygote and are *pluripotent*. Thus, they are characterized by *self-renewal*, the capacity of indefinite proliferation while maintaining their cellular identity. Moreover, these cells are able to differentiate into a broad range of cell types derived from the three germ layers, i.e. ectoderm, mesoderm and endoderm. This can experimentally be assessed by *teratoma formation*. The ultimate test of pluripotency, which is only applicable to animal models, is the contribution to chimera formation whereby the stem cells are introduced and subsequently colonize the pre-implantation embryo followed by full development of a viable organism. Thus, this test not only assesses the stem cell developmental potential but also indicates the ability to contribute to germline transmission. The implantation of the embryo occurs at E4.5 followed by the post-implantation stage at day E5.5. Afterwards, the embryo initiates a complex transformation process whereby the cells from the inner cell mass gradually lower their *stem cell potency* and initiate a highly coordinated cell specialization process. As a result, the initial pluripotent stem cell population differentiates into adult stem cells, i.e. become *multipotent*. Adult stem cells are typical examples of multipotent stem cells. After this stage, *oligopotent* and *unipotent stem cells* are generated as a continuation of the differentiation process that culminates with terminally differentiated cells. The result of this highly specialized and hierarchical process is a viable organism. A complete overview of the embryonal development is properly presented in a literature contribution by Wobus et al.[1]

Studying embryo development *in vivo* revealed that this process occurs stepwise, whereby the cell(s) gradually lose their developmental potential passing through defined cellular states coupled with precise changes in the expression levels of key regulatory genes. Thus, being able to trap and manipulate these stages under defined conditions *in vitro* would allow fundamental understanding of mammalian development and enable potential clinical applications.

1.1.2. (Pluripotent) stem cell states

The regulatory network responsible for pluripotency is highly interconnected and encompasses the interplay of various signaling and metabolic pathways, chromatin-modifying enzymes, regulatory RNA molecules, transcription factors and other proteins.[2] Studies related to the involvement of transcription factors in the establishment and maintenance of the pluripotent state identified an autoregulatory network composed mainly of *Pou5f1* (hereafter referred to as *Oct4*), *Nanog* and *Sox2* that serve as the core set of components responsible for this process. These three transcription factors share binding sites within promoters of various genes including their own. The combinatorial scrambling at the transcription sites directly modulates the expression levels of entire gene clusters, finally regulating stem cell fate. For example, lineage-differentiation programs are successfully suppressed in embryonic stem cells, while keeping key pluripotency associated genes at such levels that allow their continuous propagation.[3,4] Besides their involvement in pluripotency maintenance, *Oct4, Nanog* and *Sox2* play crucial roles in the early lineage specification processes of stem cells as well.[5] Interestingly, the three factors are expressed to certain extents in all pluripotent stem cells derived to date from human and mouse, suggesting a common mechanism of sustaining pluripotency across (mammalian) species.[6] In addition to *Oct4, Nanog* and *Sox2,* other components were also characterized and shown to contribute to the regulation of pluripotency in embryonic stem cells derived from mouse such as *Nr0b1* (also known as *Dax1*),[7] *Smad1, Smad5* and *Smad8,*[8] *Tbx3* and *Klf4,*[9] *Nr5a2*[10] and *Sall4.*[11] Therefore, the machinery responsible for the establishment and maintenance of the pluripotent state is far more complex than initially assumed and insinuates the following features: (1) pluripotency can be sustained using different combinations of transcription factors; (2) compensatory effects are possible, i.e. transcription factors from related or distant families can compensate the absence of others; (3) defined expression levels of *Oct4, Sox2* and / or *Nanog* or a combination thereof allow stabilization of different pluripotent states *in vitro* that might closely mimic *in vivo* embryonic developmental stages.

Therefore, one of the long standing goals in stem cell research has been the isolation of (pluripotent) stem cells, their maintenance and propagation in cell culture. Thus, defined protocols that permit trapping different developmental stages of the embryo *in vitro* would inevitably grant a comprehensive understanding of mammalian development and will further advance our knowledge in the area of stem cell biology.

The organism of choice for studying embryonic development in vertebrates is the mouse, primarily due to the lack of stringent ethical issues concerning embryo manipulation. So far, various pluripotent stem cell lines have been characterized in culture either by direct isolation of cells from various developmental stages of the mouse embryo or by interconverting already established cell populations under defined conditions. For example, mouse embryonic stem cells (mESCs) were first derived in 1981 from the inner cell mass of the pre-implantation blastocyst.[12,13] This initial success and the following investigations facilitated the isolation and characterization of the corresponding human counterparts by Thomson et al. in 1995.[14] 25 years after the isolation of mouse embryonic stem cells, Brons[15] and Tesar[16] independently reported the generation of mouse epiblast stem cells (mEpiSCs) from the post-implantation blastocyst. Their results suggested that additional stem cell lines can be stabilized *in vitro* under carefully optimized conditions. Besides embryonic stem cells derived from human and mouse, equivalent cellular states from various other species such as rats,[17,18] hamsters,[19] rabbit,[20] chicken,[21,22] monkeys[23] and marmosets[24] were generated.

These findings prompted further searches for culture conditions that would allow the *in vitro* stabilization of cellular states that feature restricted differentiation potential. For example, embryonic germ cells (EG) have been isolated by extracting primordial germ cells (PGC)[25] from post-implantation embryo at developmental day E8.5-E12.5. The propagation under defined conditions of the PGC population generated EG cells that exhibited all pluripotency hallmarks, i.e. self-renewal, teratoma formation and germline contribution.[26,27,28] Additional pluripotent stem cell lines were successfully isolated in the following years and comprehensively characterized. For example, early primitive ectoderm-like (EPL) stem cells were generated[29] *in vitro* from embryonic stem cells under defined culture conditions. Compared to the parent embryonic stem cell population, these cells do not contribute to chimera. In-depth characterization revealed that this cell population might represent an intermediate state between embryonic and epiblast stem cells. Blastocyst-derived stem cells (bFGF, Activin and BIO-derived stem cells, termed FAB-SC) were reported in 2008 by Chou et al.[30] Deep molecular and functional studies of this cell population revealed marked differences compared to embryonic and epiblast stem cells. Additionally, although the FAB-SC cells express the common pluripotency markers such as *Oct4*, *Nanog* and *Sox2*, this cell population fails the pluripotency tests such as teratoma formation and chimera contribution. When the FAB-SC cells are stimulated with a specific growth factor cocktail, this population regained all the pluripotent features. In addition, two types of lineage-restricted stem cells have been successfully derived. Trophoblast stem cells (TS)[31] and extra-embryonic endoderm stem cells (XEN)[32] were exhaustively characterized from multiple perspectives, and have been demonstrated to only give rise to cells derived from trophectoderm and primitive endoderm, respectively.

Literature reports so far suggest that a plethora of developmental stages, derived especially from mouse, can be isolated and propagated *in vitro* under carefully optimized conditions. These valuable discoveries offer the possibility to gain detailed insights into the developmental processes *in vitro*, to reveal signaling pathways governing lineage specification and finally offer the opportunity to manipulate cell fate for potential therapeutic benefit.

1.1.3. Integrative genetic methods to access pluripotent cellular states

As mentioned in the previous section, (pluripotent) stem cell lines have been generated by using mainly two approaches. First, cells were isolated from various developmental stages of the mouse embryo and cultured *in vitro* under defined conditions. Most of these protocols were discovered by combinatorial mixing of growth factors, vitamins, amino acids, small molecules etc. in an attempt to maintain and propagate cell populations *in vitro*. The main drawback of this approach is the continuous need to sacrifice mouse embryos for experimental purposes. The second approach employs already characterized cell lines which are used as starting point to generate new cell populations by carefully manipulating culture conditions. Both strategies have proven to be successful along the years and have yielded cellular states with various differentiation potential that may indeed correspond to *in vivo* developmental stages. In contrast to the advances made in the area of mouse stem cell biology, the research of human embryonic stem cells (hESCs) has developed at a rather slow pace due to the ethical concerns related to sacrificing human embryos for experimental purposes. Thus, for a comprehensive exploration of the potential offered by stem cell biology, the development of novel protocols granting unlimited access to mouse and human embryonic stem cells without the requirement of sacrificing living embryos are urgently sought after.[33]

The groundbreaking discovery of induced pluripotent stem cells (iPSCs) has reshaped the research landscape in the area of stem cell biology. The rapid development of this concept culminated in the Nobel Prize in Medicine Award in 2012 to Sir John B. Gurdon and Shinya Yamanaka for demonstrating that *"...mature cells can be reprogrammed to become pluripotent"*.[34] Their paramount discovery in the area of cellular reprogramming has set the stage for a new era of exploring developmental biology and cellular lineage specification *in vitro*. On one hand, Sir John B. Gurdon made significant contributions to nuclear reprogramming[35,36] using frogs as model organisms. His findings were further applied to higher organisms as well, such as sheep[37] and mouse.[38] Importantly, his discoveries pointed out the capacity of a fully differentiated cell to completely reset its cellular program and revert to a state similar to embryonic stem cells, i.e. a somatic cell is able to transform to cellular states with higher differentiation potential. Because of the technical hurdles associated with the nuclear transfer procedures, researchers sought alternative routes to access such cellular states. Along those lines, Kazutoshi Takahashi and Shinya Yamanaka reported in 2006 that mouse embryonic fibroblasts (MEFs) can be reprogrammed to an embryonic stem-like state by overexpressing four transcription factors (currently known as Yamanaka factors), *Oct4*, *Klf4*, *Sox2* and *Myc*.[39] The subsequently generated cell population was termed *induced pluripotent stem cells* (iPSCs). These cells exhibited most of the embryonic stem cell characteristics such as cell morphology and shape, hypomethylation of pluripotency-associated markers such as *Nanog* and *Fbxo15* (also known as *Fbx15*) and expressed specific embryonic stem cell markers like *Fut4* (more popular known as stage-specific embryonic antigen 1, SSEA-1 or CD15) and alkaline phosphatase (AP). Moreover, the gene expression profile of the newly generated population clustered rather well with that of the embryonic stem cells than with the initial fibroblast population. In addition, the cells exhibited a normal karyotype, formed teratoma and chimerized embryos. In-depth phosphoproteomic and transcriptomic analyses have revealed only subtle differences between embryonic and induced pluripotent stem cell lines,[40] whereas epigenetic analysis by means of global methylome profiling reported hotspots of aberrant reprogramming at the DNA methylation level.[41] Further investigations have shown that continuous culture of induced pluripotent stem cells *in vitro* promotes the accumulation of genetic aberrations,[42,43] as well as the induction of an immune response in recipients that received differentiated cells produced from iPSCs.[44] As a result, better reprogramming methods that can avoid genetic aberrations and remove the epigenetic memory barriers are likely to produce induced pluripotent stem cells of high quality that closely resemble the *in vivo* counterparts, i.e. embryonic stem cells.

The first induced pluripotent stem cells reported by Yamanaka et al.[39] were not able to give rise to live-born chimeric mice and the reprogramming yields were very low. These aspects were shortly delved by the same group[45] and in the same time by other laboratories.[46,47] One year later, human somatic cells were successfully converted to human induced pluripotent stem cells (hiPSCs) by various groups using the same four transcription factors as mentioned above.[48,49] However, Yu et al.[50] showed that another combination of transcription factors comprising *Oct4*, *Sox2*, *Nanog* and *Lin28a* could also promote the derivation of pluripotent stem cells when overexpressed in human somatic cells. The continuous optimization of the transcription factor combinations currently allows improved reprogramming yields compared to the initial results reported by Yamanaka et al.[39] Moreover, recent studies have revealed that not only the transcription factor identities are crucial for a high yielding reprogramming process to occur but also the stoichiometry of the factors transduced. For example, the overexpression of *Sall4*, *Nanog*, *Esrrb*, *Lin28a* was recently reported to efficiently generate high-quality induced pluripotent stem cells than with *Oct4*, *Klf4*, *Sox2* and *Myc*.[51] These parameters seem to affect not only the quality of the generated induced pluripotent stem cells but also influence their final epigenetic state.[52] Factors such as the origin of the somatic cell apparently affect not only the reprogramming yields but also the differentiation potential of the induced pluripotent stem cell population.[53]

The current view of the cellular reprogramming process consists of a stepwise mechanism encompassing global metabolic,[54,55] proteomic[56] and epigenetic[57,58] rewiring of the cellular components and comprises multiple intermediate states characterized by specific barriers that need to be overcome.[59,60] Research in the area of cellular reprogramming will most likely allow the identification of novel components of the pluripotency network, whose modulation would grant improved reprogramming efficiencies and higher quality of the derived induced pluripotent stem cells. Although this would allow the generation of novel cell lines, the genetic intervention associated with transcription factor overexpression would hamper the application of iPSC-based technology in regenerative medical areas such tissue engineering, cell replacement therapies, organ production and others. Thus, equivalently effective non-integrative methods that avoid the genetic manipulation of the starting population are currently in high demand for exploiting the full clinical potential of stem cell research.

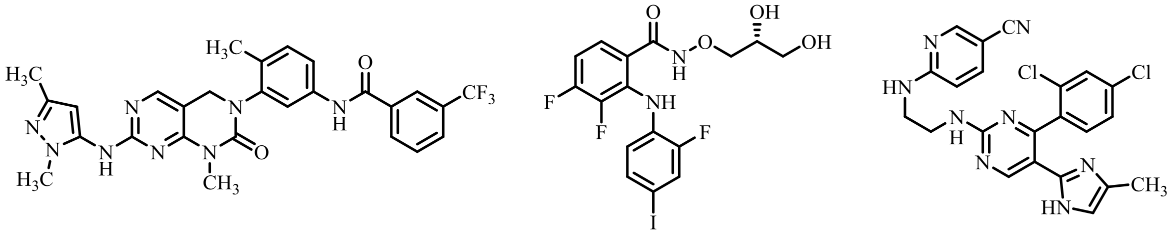
1.1.4. Non-integrative genetic methods to access pluripotent cellular states

Due to the necessity of long-lasting expression of the transcription factors in reprogramming protocols, the first induced pluripotent stem cell lines derived from mouse and human somatic cells were generated using virus-based genome integrating systems. However, the viral integration of exogenous genetic material raised numerous concerns[61] in the scientific community. Issues related to the residual expression and / or reactivation of the inserted genes, cell death and senescence, immunogenicity[44] and others still represent major hurdles that need to be addressed before considering clinical applications. In addition, some of the reprogramming factors are reported to be tumorigenic, especially *Myc*.[62] Indeed, the introduction of the transcription factor *Myc* in the reprogramming cocktails was reported to induce formation of tumor in chimeric mice.[63] Thus, in order to fully exploit stem cell technologies for clinical applications, the integration of foreign DNA needs to be avoided. For this reason, scientists oriented themselves towards other methods that can reduce or even completely eliminate transgene integrations. Recently employed protocols involve direct delivery of various types of RNAs such as miRNA (microRNA), mRNA encoding for the reprogramming factors, successive transfections with non-replicating plasmids, direct delivery of proteins, use of excisable transgene systems such as *Cre / loxP* recombination technology and addition of small molecules that can substitute for the transcription factors during the reprogramming process. A detailed description of the current technologies available for the generation of induced pluripotent stem cells including non-integrative methods can be found in an excellent contribution of Hu.[61] Moreover, a recent report shows a direct comparison among the non-integrating reprogramming methods that lead to the generation of human induced pluripotent stem cells.[64] According to this study, the most efficient non-integrative method is based on the direct delivery of mRNA encoding for the Yamanaka transcription factors. This procedure reproducibly generated reprogramming yields of approximately 2 %. Thus, advances in the field of non-integrating methods clearly made the production of transgene-free induced pluripotent stem cells possible, albeit with low yields. Future investigations hold the promise of generating large quantities of high-quality stem cells amenable for clinical applications.

1.1.5. Small molecules in stem cell research

One of the most appealing non-integrating methods to promote and maintain the pluripotent state is the use of small molecules. Indeed, these agents have emerged as elegant and safe alternatives to transgene-based methods, furnishing encouraging results in research areas such as self-renewal, differentiation and reprogramming both *in vitro* and *in vivo*. As small molecules can modulate various targets involved in signaling, epigenetic, metabolic and transcriptional processes, they represent valuable tools to probe stem cell biology and to manipulate cellular fate at will, using straightforward and defined conditions.

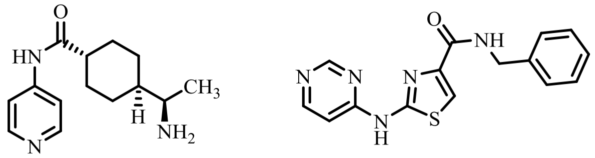
The numerous reports involving small molecule approaches to stem cell biology are not surprising considering their advantages compared to genetic integration-based methods, i.e. transgene-based methods. First, the use of small molecules does not necessitate any kind of intervention within the host genome. Second, the mode of action of small molecules is, in most of the cases, rapid and reversible leading to transient perturbation of the cellular system exemplifying a higher degree of temporal modulation. Third, small molecules can be administered in combination either sequentially or at the same time using a wide range of concentrations. A more comprehensive discussion related to the advantages of small molecules over classical genetics approaches can be found in section *1.2.2. Chemical versus genetic intervention to probe biology*. As a result of these notable features, small molecules have proven to be crucial in almost all areas related to stem cell research, such as maintenance (self-renewal), differentiation, transdifferentiation and reprogramming.



Pluripotin
RasGAP/ERK1 inhibitor

PD0325901
MEK1 inhibitor

CHIR99021
GSK3β inhibitor



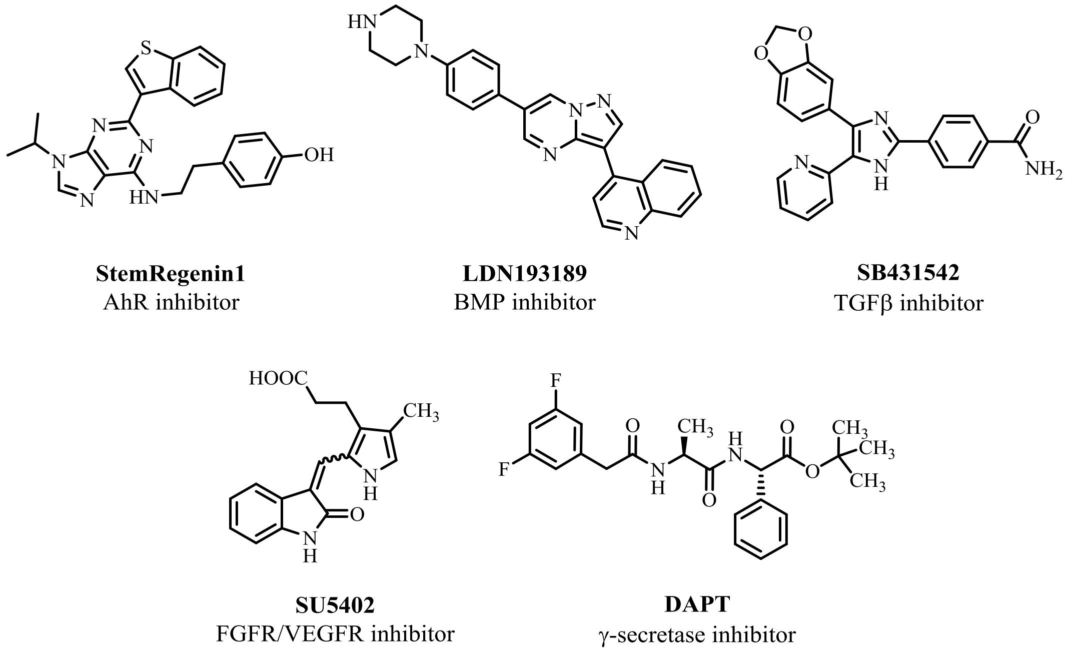
Y27632
ROCK inhibitor

Thiazovivin
ROCK inhibitor

Figure 2. Chemical structures of small molecules modulating various processes in stem cell research areas. RasGAP = Ras-GTPase activating protein; ERK1 = Extracellular signal-regulated kinase-1; MEK1 = Mitogen-activated protein kinase kinase 1; GSK3β = Glycogen synthase kinase 3 beta; ROCK = Rho-associated protein kinase.

For example, a high-throughput phenotypic screen in mESCs yielded pluripotin (see chemical structure in **Figure 2**),[65] a small molecule capable of promoting self-renewal over multiple passages in the absence of feeder cells, cytokines such as Leukemia Inhibitory Factor (LIF), Bone Morphogenetic Protein 4 (BMP4) or serum components. Deeper investigations divulged that pluripotin targets Ras-GTPase activating protein (RasGAP) and extracellular signal-regulated kinase-1 (ERK1) and maintains self-renewal via a mechanism independent of Wnt / β-catenin, LIF / STAT3 (Signal transducer and activator of transcription 3) or BMP4 / Smad. Other studies have revealed that self-renewal of pluripotent stem cells can also be achieved if the signaling pathways responsible for differentiation are inhibited upon small molecule treatment. For example, the dual inhibition of MEK1 and GSK3β with PD0325901 and CHIR99021 respectively (see chemical structures in **Figure 2**) in combination with LIF (termed 2i / LIF conditions)[66] is able to support the long-term self-renewal of embryonic stem cells derived from mouse,[67,68] rat[17,18,69,70] and even naïve human induced pluripotent stem cells[70,71] (see section *1.1.8. Differences between mESCs and mEpiSCs*). These experimental observations ultimately illustrated the high degree of conservation among signaling pathways responsible for stem cell self-renewal across species. Similar protocols using small molecules that inhibit differentiation mechanisms of adult stem cells were also developed.[72] Other examples in the area of stem cell self-renewal include Y27632 (see chemical structure in **Figure 2**),[73] a Rho-associated kinase (ROCK) inhibitor that promotes the maintenance of hESCs after cell dissociation. In a parallel study, Xu et al.[74] found in a phenotypic screen another structurally unrelated ROCK inhibitor, termed Thiazovivin (see chemical structure in **Figure 2**), that significantly increased the stem cell survival after single-cell dissociation as well as the self-renewal capacity.

A largely unexplored research area is the identification of defined culture conditions that allow the expansion of adult stem cell / tissue-specific stem cells and their progenitors *in vitro*. One of the main challenges in the area of transplantation of hematopoietic stem cells (HSCs) is the production of sufficient cell numbers for treatment of blood disorders such as leukemia. Recently, a high-content chemical screening campaign yielded StemRegenin1 (see chemical structure in **Figure 3**),[75] a purine derivative that promotes HSC self-renewal *in vitro* via the modulation of the aryl hydrocarbon receptor (AhR).



StemRegenin1
AhR inhibitor

LDN193189
BMP inhibitor

SB431542
TGFβ inhibitor

SU5402
FGFR/VEGFR inhibitor

DAPT
γ-secretase inhibitor

Figure 3. Chemical structures of small molecules modulating processes in various stem cell research areas. AhR = Aryl hydrocarbon receptor; BMP = Bone morphogenic protein; TGFβ = Transforming growth factor beta receptor 1; FGFR = Fibroblast growth factor receptor; VEGFR = Vascular endothelial growth factor receptor.

Another extensively explored research area encompasses the use of small organic molecules to promote *in vitro* differentiation of pluripotent stem cells into various cell lineages. In an early study, Chambers et al.[76] showed that a combination of LDN193189 (Bone morphogenic protein signaling pathway inhibitor) and SB431542 (TGFβ inhibitor) (see chemical structures in **Figure 3**) promotes rapid and efficient (more than 80 %) induction towards neural lineage in monolayer culture. An improved protocol was later reported, whereby the addition of three small molecules to the above mentioned cocktail, i.e. CHIR99021, SU5402 (FGFR and VEGFR inhibitor), DAPT (γ-secretase inhibitor) (see chemical structures in **Figure 3**) induced differentiation of human pluripotent stem cells toward nociceptor neurons with high efficiency (more than 75 %).[77] Further research identified small molecule combinations that promote the stem cell derivation towards endoderm,[78] cardiac[79,80] and ectoderm lineages.[81]

One of the main research areas that hugely benefited from the advent of small molecules is the cellular reprogramming of somatic or partially differentiated cells to embryonic stem-like cells.

1.1.6. Chemical reprogramming

Chemical reprogramming is a robust and viable alternative for transgene-free mediated conversion of somatic cells to pluripotent stem cells. Numerous studies showed the fascinating capacity of small molecules not only to replace crucial components of the pluripotency network, but also to increase the reprogramming yields with or without genetic intervention. These include small molecules targeting epigenetic modifiers such as DNA methyl transferases (DNMTs), histone deacetylases (HDACs) and components of signaling pathways such as kinases, enzymes involved in cell metabolism and other biomacromolecules.

For example, early reports showed that the 5-azacytidine (see chemical structure in **Figure 4**), a DNA methyl transferase inhibitor, is able to increase the reprogramming yield when somatic cells were transduced using Yamanaka factors.[82] Moreover, this compound is able to convert even an intermediate reprogrammed cellular state to embryonic stem-like cells, suggesting the major involvement of epigenetic modifiers in resetting global epigenetic marks.[82] Increased reprogramming efficiency[83,84] was also reported with popular HDAC inhibitors such as Trichostatin A (TSA), vorinostat (SAHA) and sodium butyrate (NaB) (see chemical structures in **Figure 4**). Even a structurally simple small molecule such as Vitamin C (see chemical structure in **Figure 4**) was shown to boost the reprogramming efficiencies in various contexts,[85,86,87] revealing the importance of Vitamin C-dependent histone demethylases during cellular transition processes.

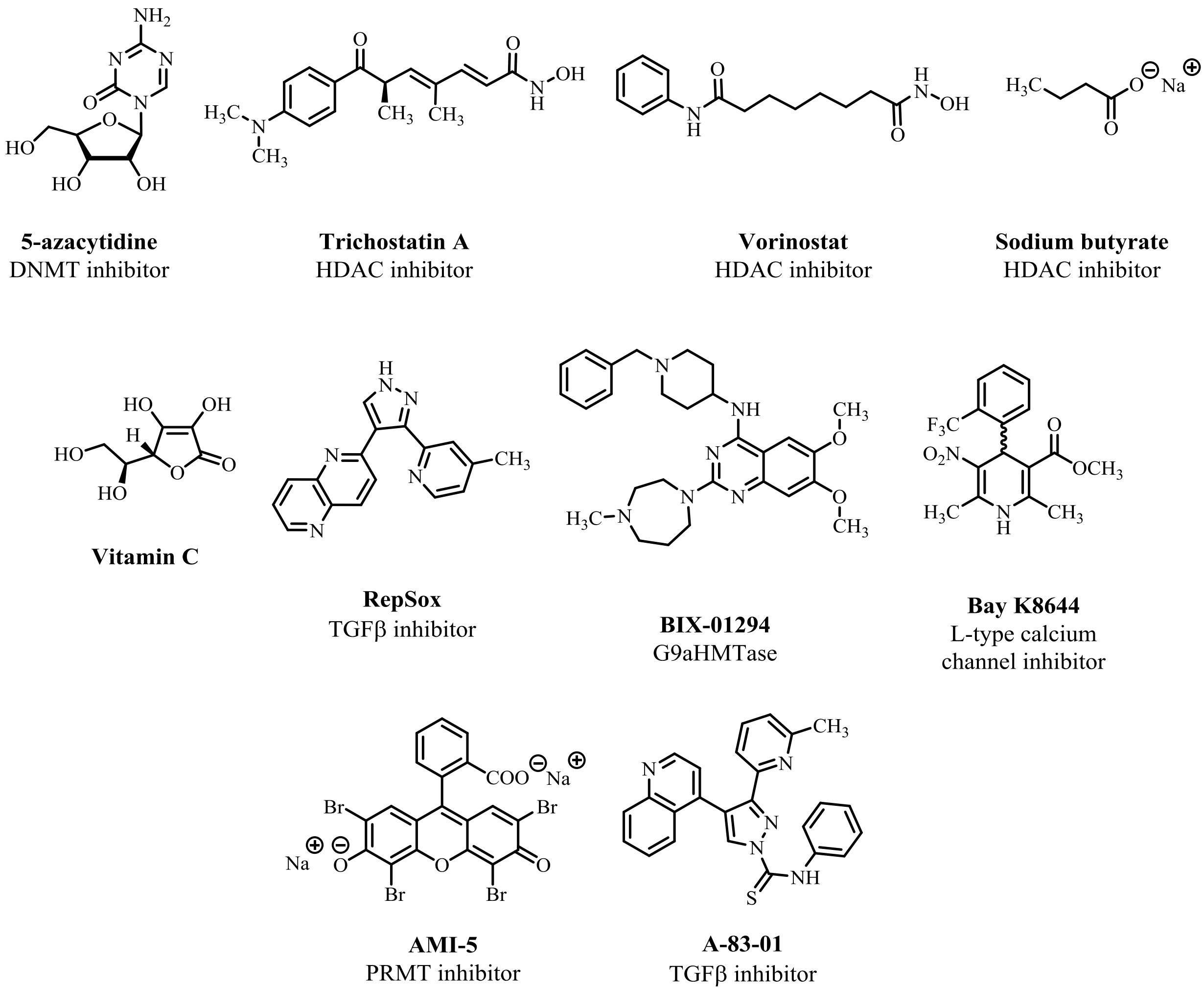

5-azacytidine
DNMT inhibitor
Trichostatin A
HDAC inhibitor
Vorinostat
HDAC inhibitor
Sodium butyrate
HDAC inhibitor
Vitamin C
RepSox
TGFβ inhibitor
BIX-01294
G9aHMTase
Bay K8644
L-type calcium channel inhibitor
AMI-5
PRMT inhibitor
A-83-01
TGFβ inhibitor

Figure 4. Chemical structures of small molecules modulating processes in various stem cell research areas. DNMT = DNA methyl transferase; HDAC = Histone deacetylase; TGFβ = Transforming growth factor beta receptor 1; G9aHMTase = G9a histone methyltransferase; PRMT = Protein arginine methyltransferases.

These initial successful studies prompted researchers to reduce the number of overexpressed transcription factors required during the cellular reprogramming and replace by small molecules. Thus, a small molecule inhibitor of TGFβ termed RepSox (see chemical structure in **Figure 4**) was reported to act as a surrogate for *Sox2* when somatic cells were transduced with *Oct4*, *Klf4* and *Myc*.[88] Even in the absence of *Myc*, the compound was still active and therefore reduced the number of transcription factors to just *Oct4* and *Klf4*.[89] Similarly, a cocktail of Bay K8644 (L-type calcium channel inhibitor) and Bix-01294 (G9a histone methyltransferase inhibitor) (see chemical structures in **Figure 4**) facilitated the reprogramming to pluripotent stem cells by ectopically expressing *Oct4* and *Klf4* in somatic cells.[90] In 2011, Ding et al.[91] proved that reprogramming is possible with the overexpression of *Oct4* alone. Thus, the Yamanaka factors were reduced to just *Oct4* supplemented with a cocktail of small molecules comprising AMI-5, a protein arginine methyltransferase (PRMT) inhibitor and A-83-01 (TGFβ inhibitor) (see chemical structures in **Figure 4**). A comprehensive description of small molecule modulators of the reprogramming process can be found in excellent review contributions listed here.[92,93]

Due to these findings, now a toolbox of compounds is available that is expediting the dissection of the biological complexity behind cellular reprogramming. This recently culminated with the generation of chemically induced pluripotent stem cells (ciPSCs),[94] a transgene-free reprogramming method of mouse somatic cells that employs a cocktail of seven small molecules. This study emphasized that chemical reprogramming is feasible even from a fully differentiated state, when the right dosage and combination of perturbing agents are applied. However, the low conversion efficiencies (approximately 0.2 %) are still not optimal for clinical applications. Nevertheless, continuous exploration of the chemical space[95,96,97] for finding active ingredients acting on cellular reprogramming may likely address the low yields and stem cell heterogeneity in future.

1.1.7. Potential applications of stem cell technologies

Stem cell research has substantial influence on the future medicine despite many challenges that lie ahead[98] and the slow progress.[99] Regenerative medicine,[100,101] which includes cell-based therapies for various diseases,[102,103] tissue engineering[104,105] and disease modeling[106,107] can directly benefit from advances in stem cell biology, pushing even further the impact of potential applications that society can benefit from.

The possibility to access any cell type of the human body via defined differentiation protocols from human pluripotent stem cells has a major impact in regenerative medicine research areas such as tissue engineering, cell replacement therapies and regeneration of failed organs. The aim of the regenerative medicine is to restore the normal function of the human cells via replacement. As a consequence, the main starting point for regenerative medical therapies is the large-scale production of various types of human cells. At present, the most appealing cell types to start with are human pluripotent stem cells but adult stem cells and somatic cells are an important part of regenerative medicine as well. One of the main reason for using human pluripotent stem cells lies on their ability to differentiate into all cells of a human body such as hepatocytes,[108] β cells,[109] cardiomyocytes[110] and cells derived from neural lineages such as motoneurons.[111] As previously mentioned, the safety issues related to the use of transgenes are still a major concern.

More promising results in the area of regenerative medicine are supplied by the multipotent adult stem cells. Various pools of adult stem cells have been characterized *in vivo,* for example in the hematopoietic system (hematopoietic stem cells), liver (liver stem cells), gut (intestinal crypt stem cells) and others. Although the adult stem cells have slower proliferation rate and lower differentiation potential *in vitro* than pluripotent stem cells, they offer a diminished risk of tumor initiation after transplantation, as recently reported.[112] As a result, there are more than 1000 open clinical trials using hematopoietic stem cells registered at the National Institute of Health (NIH).[92] Despite human induced pluripotent stem cells were first generated in 1998, the US Food and Drug Administration (FDA) agency approved only five stem cell-based medical products up to date.[99] These are exclusively derived from hematopoietic stem cells and used to cure blood and immunological disorders. Several reasons are held accountable for such scarcity of stem cell-based therapies. One of the most important factors is the necessity of large-scale production of human pluripotent stem cells and their expansion or differentiation under well-defined (xeno-free, if possible) conditions. For example, if applied in the clinical context, cardiac cell-based therapies would require approximately 10^8-10^{10} cells per patient[113] since a myocardial infarct damages approximately 10^9 myocytes as previously reported.[114] Other obstacles related to high costs, patient recruitment, limited funding, regulatory stringency and public opinion contribute to the relatively slow development of clinically relevant stem cell products.

Another appealing area of applying stem cell research is the generation of functional organs.[115,116,117] For instance, combining the capacity of pluripotent stem cells to differentiate in any cell type with the scaffolding ability of different polymers or biomaterials, led to the assembly of various functional tissue constructs which are readily available for transplantation. Along the same lines, a recent study illustrated the production of a functional human liver by transplanting liver buds obtained by *in vitro* differentiation protocols of human pluripotent stem cells.[118]

Disease modelling is another research area that would profit from advances in the generation of human pluripotent stem cells. Thus, it is nowadays possible to collect somatic cells from a patient, such as fibroblasts. Following defined reprogramming protocols, these differentiated cells can then be converted to pluripotent stem cells. Upon optimized conditions these cells can be induced to differentiate into disease-relevant cell types which usually recapitulate the disease phenotype. Then, drug screens with clinical candidates can be performed at this stage to assess the clinical efficacy in a patient-oriented manner. The ability of the pluripotent stem cells to differentiate into all types of cells of a human body enable unprecedented possibilities to evaluate clinical candidates, perform phenotypic screens and asses cytotoxicity in an patient-oriented fashion. These features represent decisive steps towards personalized medicine that overcome the limited accessibility to primary cells isolated from patients.

Most of the advances in the stem cell research were possible due to continuous efforts to decipher the signaling pathways governing various processes related to stem cell biology such as differentiation, self-renewal and reprogramming. Further investigations will most likely illuminate important discoveries in the field that might bring stem cell therapies one step closer towards the clinical applications.

1.1.8. Differences between mESCs and mEpiSCs

The seminal paper published by Nichols and Smith[119] introduced two phases of pluripotency, namely naïve and primed pluripotency, to better distinguish the various pluripotent cellular states isolated and characterized so far. This concept can best be exemplified using mESCs and mEpiSCs, which define naïve and primed pluripotent cellular states respectively. A comprehensive comparison between these two pluripotent states can be found in recent literature contributions as well as in **Table 1**.[6,120]

For example, whereas mESCs grow as small and oval domed-shape three-dimensional colonies, mEpiSCs form large and round flattened structures *in vitro* and barely survive clonal propagation from single cells in culture. Moreover, whereas mESCs are maintained *in vitro* in the presence LIF / STAT3 and BMP / Smad, mEpiSCs are propagated in growth media supplemented with FGF / Activin A. Both mESCs and mEpiSCs are pluripotent, i.e. they exhibit indefinite self-renewal and can differentiate into cells derived from the three germ layers. However, in contrast to mESCs, mEpiSCs contribute poorly to chimera formation. Thus, mEpiSCs lose their capacity to colonize the inner cell mass of an embryo at the blastocyst stage, i.e. their ability to participate in germline transmission is compromised.

	Pluripotency	
Features	**Naïve**	**Primed**
Cell morphology	small and oval domed-shape	large and round flattened
Self-renewal	YES	YES
Teratoma formation	YES	YES
Chimera contribution	YES	NO
Pluripotency factors	*Oct4*, *Nanog*, *Sox2*, *Klf2*, *Klf4*	*Oct4*, *Nanog*, *Sox2*
Naïve markers	*Zpf42*, *Fgf4*, *Gbx2*, *Dppa3*, *Nr0b1*	absent
Differentiation Markers	absent	*Fgf5*, *T*
X chromosome status	XaXa	XaXi
Doubling time	14 h	18 h
Alkaline phosphatase	YES	NO
Response to 2i	self-renewal	differentiation / death
Response to LIF / STAT3	self-renewal	none
Response to FGF / ERK	differentiation	self-renewal

Table 1. Cellular and functional properties of mESCs and mEpiSCs. Xa = X-chromosome activation; Xi = X-chromosome inactivation, 2i = chemical cocktail comprising PD032590 (MEK1 inhibitor) and CHIR99021 (GSK3β inhibitor); LIF = Leukemia inhibitory factor; STAT3 = Signal transducer and activator of transcription 3; FGF = Fibroblast growth factor; ERK = Extracellular-signal regulated kinase (adapted after **Table 1** in Nichols et al. Cell Stem Cell **2009**, *4*, 487-492).

Furthermore, the doubling time of 18 h in the case of mEpiSCs compared to only 14 h for mESCs is consistent with the unusually shorter G1 phase of the cell cycle. In addition, in the case of female cells, X-chromosome inactivation (XaXi) occurs at the epiblast stage in comparison to mESCs which exhibit two active X-chromosomes (XaXa). Although mESCs and mEpiSCs express similar set of pluripotency-associated genes such as *Oct4*, *Nanog* and *Sox2*, mEpiSCs lack the expression of embryonic stem cell-specific genes such as *Klf4*, *Dppa3*, *Zpf42* (commonly known as *Rex1)* and *Gbx2*. Moreover, mEpiSCs show increased levels of the early differentiation markers such as *Fgf5*, *T* (known as *Bra*) and do not express alkaline phosphatase compared to mESCs. These discrepancies suggest that mESC and mEpiSC populations represent two distinct pluripotent stages of the embryo development: mESC step resembles the pre-implantation embryo, i.e. inner cell mass stage, whereas mEpiSC step corresponds to the post-implantation embryo (epiblast) stage. As a result, the cellular programs and signaling pathways characterizing these two cellular states are very different. On one hand, FGF / Erk pathway inhibition promotes differentiation of the naïve mESCs, whilst under these condition mEpiSCs undergo self-renewal. On the other hand, the later property is maintained in mESCs under 2i conditions while mEpiSCs differentiate. Thus, the concept of naïve and primed pluripotency greatly facilitates the characterization of the cellular states and introduces specific features to better describe pluripotent cellular states.

Surprisingly, human embryonic stem cells (hESCs) behave similarly to mEpiSCs in many aspects such as culture conditions, cell morphology and others, suggesting that the human pluripotent stem cells might resemble a similar developmental stage, i.e. hESCs feature primed pluripotency. Recently mESC-like states for human pluripotent stem cells were reported by various labs,[121,122,123] proving that naïve pluripotency can be achieved in human induced pluripotent stem cells under defined conditions. Moreover, it confirmed the previous assumptions that human embryonic stem cells represent a distinct developmental stage compared to mouse embryonic stem cells. Due to the similarities between hESCs and mEpiSCs mentioned above, the latter might be used in principle as substitute model for hESCs studies, therefore avoiding ethical issues.

Researchers possess nowadays a broad range of (pluripotent) stem cells derived from various developmental stages of rodents and human, allowing a deeper exploration of the embryo development process *in vitro*. This has a tremendous impact on the ability to model early mammalian events in a dish for the use in clinical applications for example.

1.1.9. The mEpiSC population is heterogeneous

It is nowadays well accepted that stem cell populations are heterogeneous.[124,125] As a consequence, it is not surprising that transcription factors responsible for the maintenance of pluripotency are also heterogeneously expressed both *in vitro* and *in vivo*. For example, only approximately 80 % of the mESC population expresses *Nanog*.[126,127] A similar observation was made in the case of *Zpf42,*[128] a naïve ESC-specific marker,[128,129] whose levels fluctuate in the mouse embryonic stem cell population.[130] These studies indicated that the heterogeneity observed is not due to coexistence of isolated cell populations, but rather to different populations of cells which are able to interconvert into one another. In the case of the heterogeneous *Nanog* subpopulations, it was shown that *Nanog*-positive cells maintained undifferentiated colonies for prolonged time, whereas *Nanog*-negative cells underwent differentiation under continuous propagation in culture.[126] This observation indicates that the two cell states represent slightly different developmental stages that are present within the same population. Of note, transcription factor heterogeneities were previously evidenced *in vivo* in the case of *Nanog* and *Gata6*.[131]

Following the same line of evidence, Han et al.[132] showed that mEpiSCs are analogously dissimilar, comprising a *Oct4*-positive population (>99 % of the cell population) and a *Oct4*-negative population (less than 1 % from the total amount of cells). These subpopulations are not embryonic stem cells as demonstrated by gene expression profiling, epigenetic marks and *Oct4* promoter usage. Similar to embryonic stem cells, the authors could show that the two populations are interconvertible *in vitro* and represent distinct states derived from the epiblast trapped at different developmental stages. The study could also clarify the previous observation made by Garner et al.,[133] whereby primitive ectoderm (PrE) cells isolated from E5 embryos were capable of contributing to chimera formation in comparison to cells isolated from E6-E7 embryo. By comparing the gene expression profiles of the *in vivo* and *in vitro* mEpiSCs samples, the study concluded that the *Oct4*-positive population might represent an early-stage epiblast (E5.5), whereas the *Oct4*-negative population a late-stage epiblast (E6.5). Thus, the *Oct4*-negative cell population corresponds to the most differentiated state that can be trapped *in vitro* within the epiblast stage. This study highlights the paramount importance of comprehensive characterization of the cellular heterogeneity across stem cell populations.

Mouse epiblast stem cells represent an important cellular model to study germ layer specification, because the epiblast stage is the last pluripotent cellular state before gastrulation. Recently, mEpiSCs were used as starting point to generate neural stem cells,[134] a multipotent stem cell population with the ability to differentiate into neurons and glial cells as previously reported.[135]

1.1.10. Reprogramming mEpiSCs to mES-like cells – state of the art

Various methods have been reported so far detailing the conversion of mEpiSCs to embryonic stem-like cells. For example, a simple change of the FGF / Activin A growth medium to LIF / STAT3 alone, was sufficient to promote conversion of epiblast stem cells isolated from E5.5-E7.5 mouse embryos to naïve pluripotency.[136,137] Under these conditions, dramatic changes at the genetic and epigenetic level of the starting population such as global DNA demethylation, reactivation of the X-chromosome and expression of E-cadherin were observed. Moreover, the obtained reprogrammed embryonic stem cells exhibited all pluripotency hallmarks such as self-renewal, teratoma formation and chimera contribution.[136] Other reports demonstrated that genetic manipulation of various transcription factors at the epiblast stage such as *Prdm14-Klf2*,[138] *Nr5a2*,[139] *Klf4*[140] can generate cell populations that exhibit naïve pluripotency features. Moreover, the overexpression of E-cadherin in the presence of LIF increased the transition efficiency from mEpiSCs to chimera-competent mESCs.[141] The authors revealed by genetic and chemical means that the sequestration of β-catenin inside the nucleus is responsible for the enhanced reprogramming efficiency. Recently, Zhou et al.[142] reported the conversion towards an ESC-like state by chemical means only, using a cocktail of five small molecules that include CHIR99021 (GSK3β inhibitor), PD0325901 (MEK1 inhibitor), A-83-01 (TGFβ inhibitor), PD173074 (FGF inhibitor) and Parnate, a Lysine-specific histone demethylase 1 (LSD1) inhibitor. Additionally, Rao et al.[143] could achieve the same transition using a similar protocol comprising a combination of fibroblast growth factor 2 (FGF2) and four small molecule inhibitors, namely Y27632 (ROCK inhibitor), PD032590 (MEK1 inhibitor), SB431542 (TGFβ inhibitor) and CHIR99021 (GSK3β inhibitor).

However, in all reports mentioned above, unsorted mEpiSCs were employed. As previously mentioned by Han et al.,[132] mEpiSC population is heterogeneous, consisting of two distinct subpopulations characterized by Oct4-GFP expression. The Oct4-GFP-positive cells represent less than 1 % of the entire cell population and—as reported—might correspond to the early-stage epiblast. On one hand, it has been shown that this tiny fraction has a greater developmental potential and can be converted to an ESC-like state solely by the use of 2i / LIF conditions. On the other hand, the Oct4-GFP-negative fraction represents more than 99 % of the mEpiSC population and might represent the late-stage epiblast (thereafter late-stage mEpiSCs). Surprisingly, these cells are refractive towards conversion to mESCs under 2i / LIF conditions suggesting that the late-stage population represents a distinct and less potent cellular state than the Oct4-GFP-positive population. Based on these findings, it can be surmised that the previous studies might only succeeded to reprogram the Oct4-GFP-positive cells, since an unsorted mEpiSC starting population has been used. In addition, Berneman et al.[144] and Hanna et al.[131] independently showed that epiblast stem cell lines are disparate among each other, namely some are refractive towards 2i / LIF treatment, whereas others smoothly undergo reprogramming towards an ESC-like state. This behavior was explained by the expression levels of the mesendodermal marker *T*,[145] a transcription factor part of the Tbox family involved in the formation of the mesoderm.[146,147] There is an inverse reciprocal relationship between the expression levels of this marker and the efficiency of generating ESC-like cells under stringent mESC conditions, i.e. 2i / LIF. Moreover, the overexpression of *T* in *T*-negative cells blocks the conversion capacity of these cells. This observation suggests that some mEpiSC lines are more prone to transition to higher cellular states, i.e. they possess lower reprogramming barriers than others.

1.1.11. Reprogramming late-stage mEpiSCs to mES-like cells – preliminary results

A cell-based assay in a 96-well format was performed by Dr. Damir J. Illich (Max Planck Institute of Biomolecular Medicine, Münster, Germany) under previously described conditions.[148,149] This assay monitors the reactivation of Oct4-GFP expression upon treatment of late-stage mEpiSCs with small molecules and quantifies the GFP fluorescence intensity by means of fluorescence activated cell sorting (FACS). Screening the Library of Pharmacologically Active Compounds (LOPAC®) yielded the compound 6-phenylpteridine-2,4,7-triamine (Triamterene, thereafter **TR**; see chemical structure in **Figure 5**), a known inhibitor of the epithelial sodium channel (ENaC).[150]

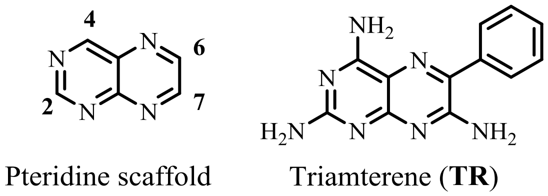


Figure 5. Chemical structure of the pteridine scaffold and Triamterene (**TR**).

In-depth investigations revealed the ability of **TR** to reprogram the initial cell population first to a pre-ESC state that features most of the embryonic stem cell hallmarks such as cell morphology, propagation conditions *in vitro*, and differentiation potential, however fails to contribute to chimera formation.[148,149] Global gene expression and epigenetic profiling studies confirmed that the generated cells are trapped in an intermediate state between mEpiSCs and mESCs.[148,149] The transition to a naïve state was only possible by combining **TR** with PD0325901 (thereafter **PD,** see chemical structure in **Figure 2,** section *1.1.5. Small molecules in stem cell research*), a known MEK1 inhibitor. This compound promotes and sustains naïve pluripotency in various contexts.[151] Under these conditions, the **TR** / **PD** combination successfully yielded pluripotent stem cells that regained the germline contribution ability.

1.2. Chemical Genetics

1.2.1. Definition

One of the major goals of natural sciences is to investigate and understand biological systems. For example, the relationship between genetic material represented by genes and their involvement in particular processes is fundamental to understand the regulation of biological systems. This correlation can be studied by means of classical genetics methods that investigate the relationship between genes and physiological effects (phenotypes) using two approaches (see **Figure 6**).

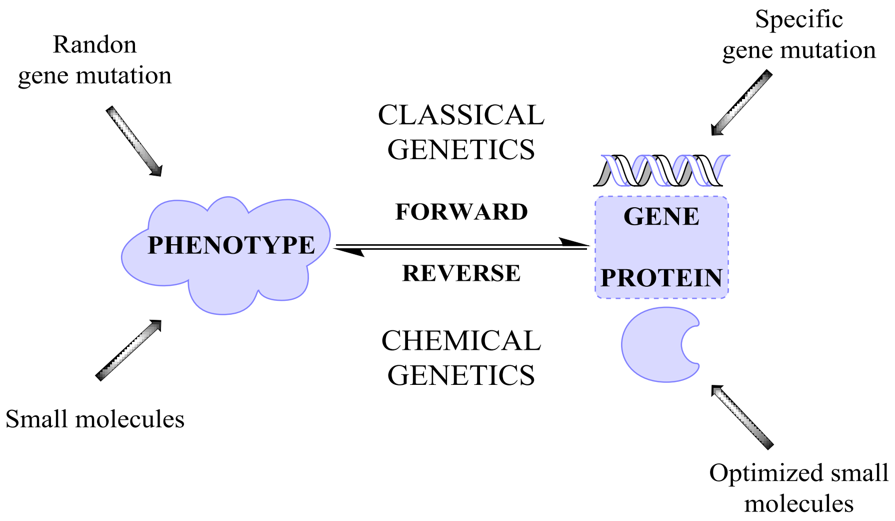


Figure 6. Concept representation for the forward / reverse genetics and chemical genetics. Forward genetics / chemical genetics involve top-down approaches and try to identify the gene / gene product responsible for a certain cellular event, i.e. phenotype. Reverse genetics / chemical genetics use a bottom-up approach where the impact of targeted gene / gene product inactivation on the phenotype is studied (adapted after **Figure 1** in Kawasumi et al. J. Invest. Dermatol **2007**,*127*(7),1577-1584).

On one hand, forward genetics allows a top-down approach whereby genes involved in a biological process of interest are identified by means of analyzing their impact on the phenotype by generating random mutations within the genome. On the other hand, reverse genetics is a bottom-up approach whereby an already selected gene is randomly mutated or the expression levels are altered. Thereafter, the impact of these modifications is used to investigate the effect on the phenotype. Both approaches furnished valuable insights into fundamental biological processes and allowed a deeper understanding of the basic principles governing living systems. These were only possible due to development of powerful molecular and cell biology techniques such as mutagenesis and manipulation of gene expression levels by various means, including gene knock-in / -out, RNA interference (RNAi) technology and other related methods.

By analogy to classical genetics, chemical genetics[152,153] explores gene products by chemical means (see **Figure 6**). In this area of research, biological systems, processes, phenomena are analyzed by means of small molecule intervention in a complementary fashion to the classical genetics methods.[154,155,156] Accordingly, forward chemical genetics approaches biological processes from phenotype (effect) to genotype (cause) using small molecules. For example, by performing a phenotypic screen with compound libraries the capacity of a small molecule to induce a certain cellular effect is determined. Biologically active compounds are identified, confirmed by complementary assays and afterwards the addressed biomacromolecule (protein, RNA or DNA or others) is deconvoluted by means of target identification techniques. Instead, reverse chemical genetics starts with a defined target, and small molecules that specifically modulate its function are revealed by target-based screening approaches. Once identified, these derivatives are employed in relevant cellular models or even in whole organisms to probe the involvement of the biomacromolecular target in the biological process of interest.

Both approaches have yielded numerous small molecules that can alter the function of various biomacromolecules such as enzymes, receptors, ion channels, RNA, DNA and others. Although the number of biological targets modulated by small molecule means is steadily increasing, this might still represent a tiny fraction of the total number of possible druggable biomacromolecules found in the human genome, i.e. human druggable genome.[157] Thus, chemical genetics intends to expand this small molecule toolbox in the attempt to furnish modulators for novel biological targets and pathways to unveil biological principles that govern living systems.

1.2.2. Chemical versus genetic intervention to probe biology

Both classical and chemical genetics approaches represent powerful tools for the study of biological processes. The advantages and disadvantages related to both methods are analyzed in the section below.

One of the main benefits of using chemical compounds over genetic intervention is the ability of a small molecule to disrupt a single module within a multifunctional biomacromolecule. Moreover, in the case when the gene of interest encodes a protein with multiple domains, chemical genetics can render specific small molecule binders for each module and thereby prevent the abolishment of the function of the entire biomacromolecule. For example, kinase inhibitors block only the phosphotransferase activity concomitantly leaving intact the scaffolding function of the protein, where biological partners might dock. Thus, small molecule intervention can dissect the two protein functions as compared with the genetic ablation which would have removed the whole protein upon gene deletion. An interesting example is tubacin,[158] a small molecule designed to only interfere with the tubulin deacetylase activity of the multifunctional histone deacetylase 6 (HDAC6) protein. Thus, tubacin just inhibits the catalytic domain that is responsible for tubulin deacetylation without perturbing the HDAC6 activity. The ability of tubacin to distinguish between histone and tubulin deacetylase activity within the same enzyme was subsequently employed in further biological studies related to the involvement of HDAC6 activity in cell motility and protein degradation.[159]

Another unique feature of small molecules is the ability to disrupt protein-protein interactions.[160,161] Therefore, this aspect further expands the applicability area of small molecules in perturbing biological systems.

Small molecules are efficient tools, especially in cases where classical genetics approaches are not effective or applicable. Although the manipulation of genomes can easily be performed in prokaryotes, viruses and bacteria, mammalian cells are usually not straightforwardly accessible to genetic intervention due to the diploid nature of their genome. However, the newest developments in genetic engineering methods such as Clustered Regularly Interspaced Short Palindromic Repeats CRISPR / Cas9[162,163] allow the generation of precise knock-in / -outs in mammalian cells relatively straightforward. Nevertheless, it is often difficult to distinguish if the observed phenotype is a result of losing the function or the scaffolding feature of the biological target when complete deletion of the gene construct occurs. In contrast, small molecules can be readily applied regardless of the biological context, either to cells or whole organisms across species. However, compounds can exhibit cytotoxicity and therefore optimization of the chemical structure occurs before probing a certain phenotype or before administration to patients in clinical trials. Chemical genetics is also employed when removal of an essential gene by genetic means would translate to lethality. In this case, working with small molecules at concentrations below the lethal dose could render a partial knock-out phenotype that might allow studying the biological significance of that specific target.

One of the most important features of using small molecules is their rapid and often reversible action in biological context. Compounds can be straightforwardly added and removed from the culture media providing the possibility to precisely manipulate the exposure time. In addition, by applying different concentrations of the small molecule one can generate dose-dependent studies of the phenotype of interest both *in vitro* and *in vivo*. However, small-molecules have often disadvantages such as specificity issues when targeting closely related protein isoforms or protein families with conserved function. Lack of specificity mostly originates from identical amino acids in the active sites among the members of a protein family. Whereas genetic methods might be able to specifically address a single protein isoform of a protein subfamily, chemical genetics often fails to supply small molecules for each particular isoform of interest. Additionally, even highly optimized small molecules exhibit unintended targets ("off-targets") in cellular context, which might impair the phenotype of interest.

A new method is currently emerging that tries to compensate for the lack of specificity of the small molecules and the lack of temporal control associated with genetic methods. The so-called allele-specific chemical genetics[164] combines the specificity achieved with genetic tools and the reversible and tunable character of chemical genetics (see **Figure 7**). This approach involves the simultaneous optimization of the small molecule target protein pair by means of recombinant DNA technology and organic synthesis, respectively.[165,166] On one hand, relying on available structural information of a protein, mutants can be generated by means of site-directed mutagenesis in two different directions. First, the mutant is engineered such that it is able to engage the natural ligand that can be outcompeted with an allele-specific derivative generated from the original compound (see **Method A**, **Figure 7**). In this way, the natural ligand binding capacity is left intact and this event can be specifically disrupted with the allele-specific derivative. On the other hand, the structure of the protein of interest can be engineered to block the binding event of the natural ligand. As a result, the protein structure and overall conformation remain intact but the binding of the natural ligand is blocked. In this case, small molecules that specifically interact with this mutant protein can be synthesized (see **Method B**, **Figure 7**).

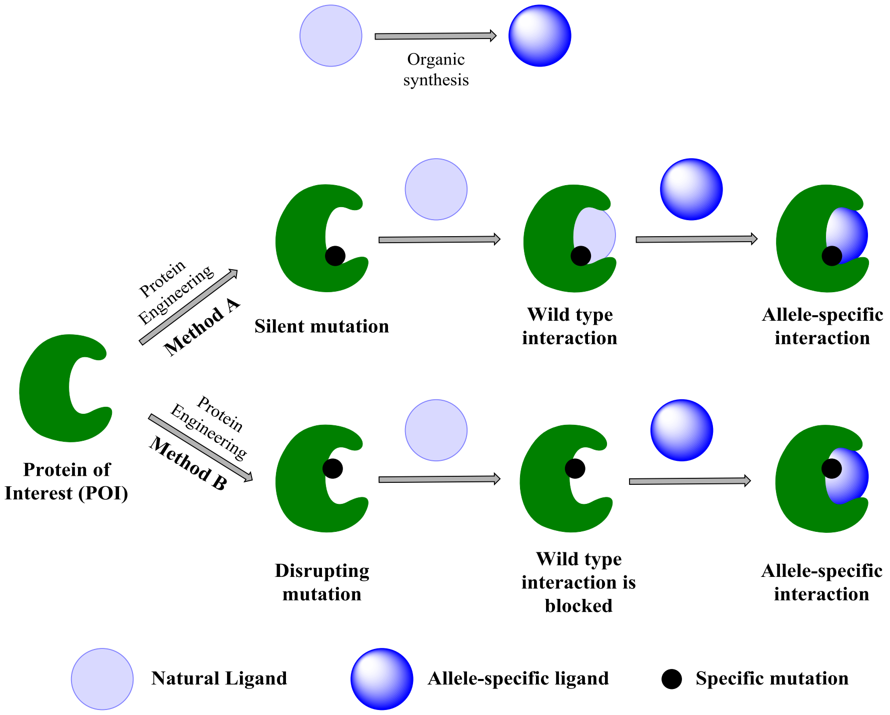

Organic synthesis
Protein Engineering
Method A
Protein Engineering
Method B
Silent mutation
Wild type interaction
Allele-specific interaction
Protein of Interest (POI)
Disrupting mutation
Wild type interaction is blocked
Allele-specific interaction
Natural Ligand
Allele-specific ligand
Specific mutation

Figure 7. Concept of allele-specific chemical genetics. Two methods are used to study the specific modulation of the protein of interest (POI) by small molecules. **Method A**: POI is specifically mutated to allow binding of the natural ligand that can be outcompeted by a chemically synthesized analogue. **Method B**: A disrupting mutation is introduced in the POI that blocks the interaction with the natural ligand while keeping the binding of the chemically synthesized analogue (adapted after **Figure 1B** from Islam ACS Chem. Biol. **2015**, *10*, 343-363).

The method has been reported to be robust and successfully implemented in the case of kinases, GTPases, methyltransferases, acetyltransferases, proteases and other classes of enzymes.[164] For example, many important observations have been made related to the differences that can arise from chemical versus genetic inhibition of protein kinases.[167] However, the allele-specific chemical genetic approach requires detailed structural information about the small molecule protein target combination in order to rationally design a complementary pair that is orthogonal to the natural variant. Additionally, the engineered protein mutant needs to conserve the natural conformation, structure and function once translated within the host proteome. Thus, the availability of such methods allows the characterization of the protein function at relevant biological time scale in a highly specific manner that would otherwise be difficult to investigate with any other technologies.

By analogy to allele-specific chemical genetics, “allele-specific RNAi” has also been developed.[168,169] This method selectively targets mutant transcripts in the context of human diseases by means of RNA interference technology, unveiling the biological impact of disease-causing mutations.

1.2.3. Forward chemical genetics applied to stem cell research

In the context of stem cell research, the forward chemical genetics approach is predominantly employed. Many examples in this area report phenotypic screens of compound libraries for studying processes such as differentiation of human pluripotent stem cells into different lineages, reprogramming of fully or partially differentiated cells or self-renewal of pluripotent stem cells. Advances in miniaturization, fluorophore development, imaging and microscopy techniques, and genetic engineering protocols nowadays enable assaying large compound libraries both faster and with higher content than ever before. Most of the chemical screens related to stem cells presented in the previous sections make use of commercial compound collections that contain small molecules or drug candidates with known targets and mode of action. For example, many compounds that act on different stem cell related processes are kinase inhibitors that have been already optimized during oncology-oriented research campaigns. As a consequence, the identification of the underlying biological mechanism is in this case relatively straightforward. For example, Aurora A kinase represents a barrier in reprogramming[170] and thus inhibition of this target by genetic or chemical means was described to facilitate reprogramming of somatic cells as revealed from a phenotypic screen monitoring the reactivation of the *Oct4*-GFP expression in somatic cells.[171] Other studies highlighted the importance of kinase inhibition in increasing the reprogramming efficiencies in various experimental setups.[171,172]

Although successful in identifying agents that manipulate stem cell fate, phenotypic screens have several drawbacks. First, depending on the phenotype under investigation, the screening throughput might be low and thus a limited number of compounds can be tested. Moreover, the price per well might hamper screening of large compound libraries also. Secondly, although a phenotypic screen can provide a plethora of information, it does not point out the cellular target(s) responsible for the biological event under investigation. At this point, various target identification and confirmation techniques are employed in order to elucidate the underlying biological target(s) addressed by the identified hit molecule(s). This process occurs stepwise and encompasses the interplay of different biological and biophysical methods rendering this process time consuming. For a comprehensive description of target identification and confirmation techniques, important review contributions are available.[173,174,175,176]

1.2.4. Chemical genetics and stem cells – future challenges

Chemical genetics principles were successfully applied in many studies related to stem cell research. However, most of these reports employed forward (chemical) genetics approaches. For example, genome-wide loss-of-function screens using RNAi or shRNA (small hairpin RNA) identified barriers in the context of cellular reprogramming.[59,60] Thus, genetic and / or chemical inhibition of the barriers allowed increased reprogramming yields for generating pluripotent stem cells. Conversely, reverse chemical genetics has not often been employed in stem cell research, although gene targets that block the reprogramming of mouse embryonic fibroblasts (MEFs) have been reported such as *Nfe2*, *Cdkn2aip* (CDKN2A-interacting protein), *Msx3*, *Dbx1*, *Lzts1* (Leucine zipper putative tumor suppressor 1), *Gtf2i* (General transcription factor II-I) and *Ankrd22* (Ankyrin repeat domain-containing protein 22).[59] Additionally, the application of a shRNA lentiviral library targeting 19.527 human genes revealed ubiquitination, endocytosis and cell adhesion / motility as roadblocks towards the generation of human iPSCs, when BJ human fibroblasts (human forseskin) were reprogrammed using OKSM transgenes in combination with *TP53* RNAi.[60] Thus, starting with a target-based screen for such barriers would be advantageous in revealing novel chemotypes that can then be evaluated in relevant stem cell-based screens. Improved organic synthesis protocols will most likely aid the synthesis of compound collections of closely related analogues that will better address the concerns of the stem cell community.

While most of the studies so far use derivatives already optimized for other research purposes such as oncology, the reverse chemical genetic approach might be able to reveal new chemical entities with improved mode of action. The knowledge gathered about the barriers in cellular reprogramming could also be used to design polypharmacological agents that engage multiple targets simultaneously. Thus, instead of using cocktails of small molecules that target two or more biomacromolecules, one could envisage the design of single agents in an attempt to reduce the complexity of the targeted interactome. Computational approaches are currently available that facilitate the *de novo* design of ligands with different pharmacological profiles.[177] Therefore, polypharmacological agents could in principle be exploited in phenotypic screens when the intended targets are addressed at the right dosage.

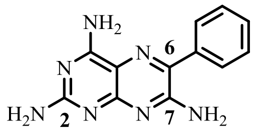
Next generation therapeutics are expected to harness the potential of endogenous stem cell pools to promote replacement of damaged cell populations *in vivo*, promote tissue repair and stimulate proliferation of dormant cells.[98,178,179] The identification of new chemicals and their addressed targets will most likely reveal the signaling pathways that govern developmental biology at the level of the whole organism.

`

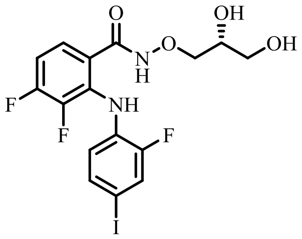
2. Aim of the Thesis

There is a continuous need for discovering new chemical entities with novel modes of action to achieve chemical reprogramming of fully or partially differentiated cells to embryonic stem-like cells. These compounds and their molecular targets would enable a better understanding of the signaling pathways governing cell dedifferentiation and would more efficiently advance stem cell-based technologies in academic and clinical settings.

The combination of Triamterene (**TR**) and the MEK1 inhibitor PD0325901 (**PD**) (see chemical structures in the figure below) has been shown to promote the reprogramming of partially differentiated cells termed late-stage mouse epiblast stem cells (mEpiSCs) to an embryonic stem (ES)-like cell population. In-depth biological studies revealed that the mechanism driving the cellular transition is mediated by casein kinase 1 (CSNK1).



Triamterene (**TR**)



PD0325901 (**PD**)
MEK1 inhibitor

Based on these preliminary results, one of the main goals of the current study is to improve the reprogramming activity of the initial hit **TR**. This can be achieved by means of synthesizing a compound collection of closely related analogs based on the **TR** scaffold bearing various substituents at positions C2, C6 and C7 (see above). Synthesis of the pteridine-based derivatives is straightforward from commercially available starting materials using methods mentioned in the literature. The access to such compounds will enable a comprehensive study of the substitution pattern(s) on the pteridine ring that allow increased cellular transformation rates of late-stage mEpiSCs to mES-like cells. The biological activity will be assessed by quantifying the re-expression of the Oct4-GFP construct, a crucial pluripotency-specific marker, in the reprogrammed cells by means of flow cytometry. In addition, the compound collection shall explore the correlation between the conversion efficiency and the extent of CSNK1 inhibition mediated by the pteridine-based derivatives. Profiling the newly synthesized derivatives for CSNK1 enzyme inhibition will be achieved by means of state-of-the-art setups that allow measuring kinase activity. This study shall reveal the structural features of the pteridine-scaffold necessary for potent and / or selective inhibition of the CSNK1 kinase family as well as the mode of inhibition with respect to ATP. Additionally, the interaction of the most promising candidates with CSNK1 isoenzymes will be analyzed in cells or cell lysates using affinity-based methods. Target engagement of ATP binding proteins, in particular kinases, by ATP competitive agents shall be demonstrated by means of activity-based profiling using the ActivX^{TM} ATP probe and immunoblotting when the target of interest is known beforehand. Since kinase inhibitors are rarely specific, special focus will be placed on the identification of further protein targets addressed by this class of compounds. This will be investigated by *in vitro* kinase profiling study and employing a chemical proteomics strategy performed in cell lysates using the ActivX^{TM} ATP probe.

3. Experimental part

3.1. Experimental part – chemistry

3.1.1. Materials

All the chemicals mentioned in the current study were purchased from Alfa Aesar and Sigma Aldrich and used as provided, without any further purification. Analytical thin-layer chromatography (TLC) was executed on silica gel aluminium plates with F-254 indicator (Merck, Germany). Compounds were visualized on the TLC plates by irradiation with UV light. Acros 60A (particle size 0.035 – 0.070) silica gel was employed for the column chromatography purification of the derivatives reported in this study. Compound characterization by means of ^{1}H- and ^{13}C-NMR analyses was performed on Bruker DRX400 (400 MHz), Bruker DRX500 (500 MHz) and INOVA500 (500 MHz). Deuterated DMSO (DMSO-d6) was used as internal standard. Signal characterization is described using the next abbreviations: s (singlet), d (doublet), dd (double doublet), dt (double triplet), t (triplet), q (quartet) and m (multiplet). Coupling constant values are presented in Hertz (Hz). Fourier transform infrared spectroscopy (FT-IR) characterizations were recorded with a Bruker Tensor 27 spectrometer (ATR, neat) and are reported in terms of frequency of absorption (cm^{-1}). Chemical yields refer to isolated and pure substances. High resolution mass (HRMS) spectra were recorded on a LTQ Orbitrap mass spectrometer coupled to an Accela HPLC-System (HPLC column: Hypersyl GOLD, 50 mm × 1 mm, 1.9µm). Solvent mixtures for thin layer chromatography are depicted as volume : volume. PE is the abbreviation for petroleum ether.

3.1.2. Methods

Derivatives **32-34** (entries **32-34**, **Table 3**) were synthesized according to literature procedures.[180] Starting materials **35** and **36** (entries **35-36**, **Scheme 3**) were synthesized following a procedure already reported in the literature.[181] Derivatives **37** and **39** (entry **37** and **39**, **Table 4**) were generated using a known synthetic route.[182] Starting materials **41-49** (entries **41-49**, **Scheme S1**) were generated starting from 4-chloro-2,6-diaminopyrimidine following a two-step procedure previously reported.[183] The confirmation of the identity and purity of the compounds mentioned in the literature was assessed by ^{1}H-NMR and HPLC-MS analyses only. Synthesis of the derivatives shown in **Table 2**, except for the compounds already mentioned above, was performed using a slightly modified workup procedure than originally reported in the literature.[182,184,185,186] The derivatives not mentioned in the literature are characterized below by means of ^{1}H- and ^{13}C-NMR, HRMS and FT-IR (see section *11. Compound characterization*).

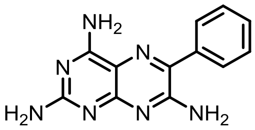
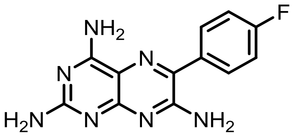
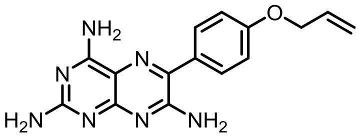
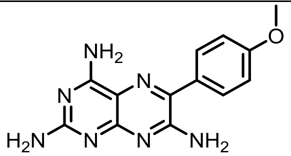
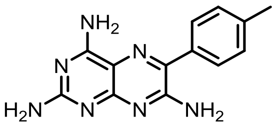
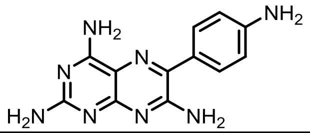
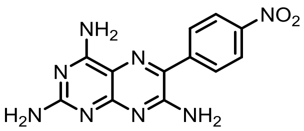
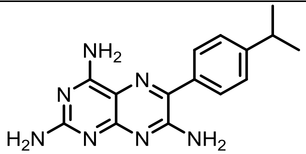
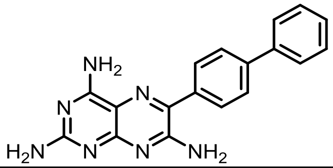
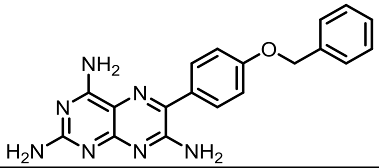
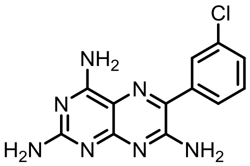
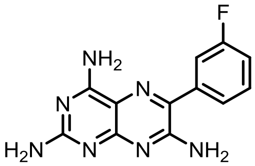
General procedure for the synthesis of derivatives **2-30** (entries **2-30**, **Table 2** and entries **60-72**, **Table A2**)

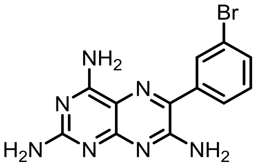
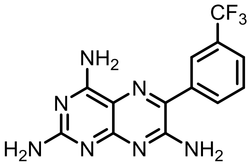
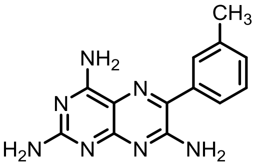
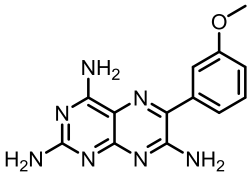
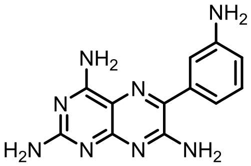
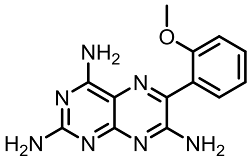
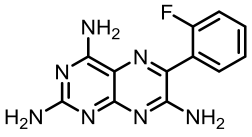
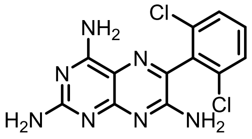
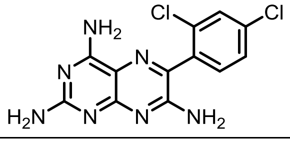
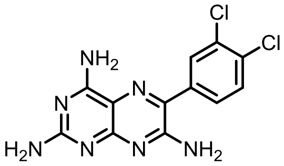
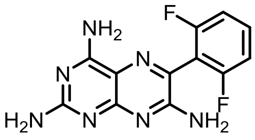
Epiblastin A (entry **12**): NaH (60 % oil dispersion, 0.75 mmol, 1.1 eq) was dissolved in 8 mL anhydrous 2-ethoxyethanol and vigorously stirred at room temperature. After H_2 evolution, a clear solution was obtained and then, (3-chlorophenyl)acetonitrile (0.75 mmol, 1.1 eq) was added followed by 5-nitroso-2,4,6-triaminopyrimidine (0.68 mmol, 1 eq). The reaction mixture was heated to 140ºC and refluxed for 2 hours. After cooling to room temperature, the mixture was concentrated *in vacuo*. The obtained solid is treated with 5 mL water, filtered and additionally washed twice with 3 mL methanol. The solid was dried to finally furnish 107 mg of a yellow solid (yield = 55 %).

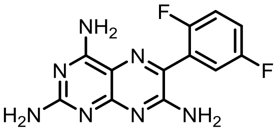
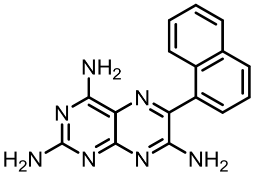
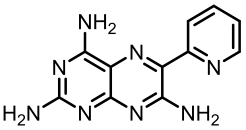
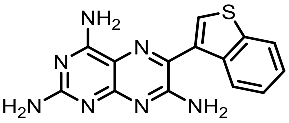
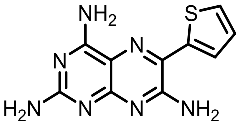
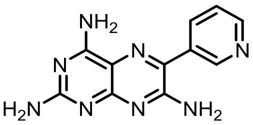
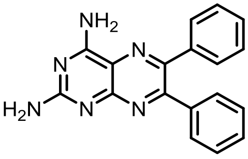
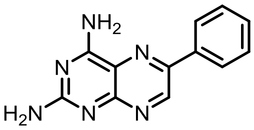
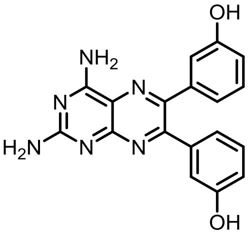
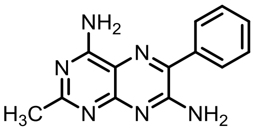
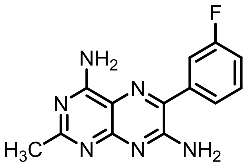
General procedure for the synthesis of derivatives **50-59 (**entries **50-59, Table 6**)

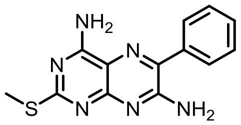
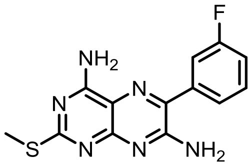
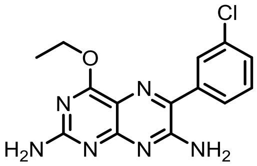
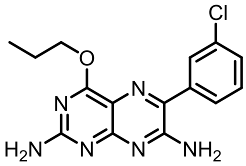
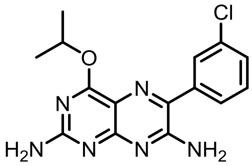
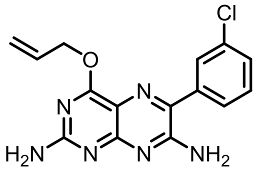
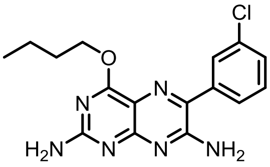
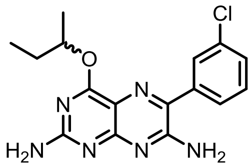
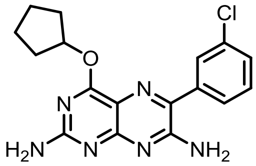
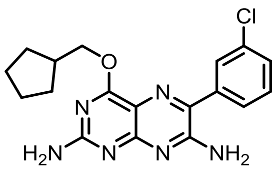
Compound **50**: (3-chlorophenyl)acetonitrile (0.5 mmole, 1.2 eq) was dissolved in 3 mL dry THF and then Cs_2CO_3 (0.5 mmole, 1.2 eq) was added followed by 4-ethoxy-5-nitroso-2,6-diaminopyrimidine (0.4 mmole, 1.0 eq). The mixture was stirred at 65ºC for 3 hours. After solvent evaporation, the mixture is purified using column chromatography on silica using acetone : petroleum ether 4:1 (R_f = 0.12) furnishing 55 mg of an yellowish brown solid (yield = 44 %).

Chemical structures and IUPAC nomenclature of the derivatives synthesized in the current study are presented in **Table A1** (see below). These compounds showed a good solubility in DMSO and stock solutions of 10 mM in 100 % DMSO could be obtained. Conversely, derivatives shown in **Table A2** were either insoluble in 100 % DMSO or precipitated out easily from the solution. As a consequence, these compounds were analyzed neither in the stem cell assay or the inhibition of CSNK1A1/D/E isoenzymes.

No.	Compound structure	IUPAC name
2		6-phenylpteridine-2,4,7-triamine
3		6-(4-fluorophenyl)pteridine-2,4,7-triamine
4		6-(4-(allyloxy)phenyl)pteridine-2,4,7-triamine
5		6-(4-methoxyphenyl)pteridine-2,4,7-triamine
6		6-(p-tolyl)pteridine-2,4,7-triamine
7		6-(4-aminophenyl)pteridine-2,4,7-triamine
8		6-(4-nitrophenyl)pteridine-2,4,7-triamine
9		6-(4-isopropylphenyl)pteridine-2,4,7-triamine
10		6-([1,1'-biphenyl]-4-yl)pteridine-2,4,7-triamine
11		6-(4-(benzyloxy)phenyl)pteridine-2,4,7-triamine
12		6-(3-chlorophenyl)pteridine-2,4,7-triamine
13		6-(3-fluorophenyl)pteridine-2,4,7-triamine

No.	Compound structure	Compound name
14		6-(3-bromophenyl)pteridine-2,4,7-triamine
15		6-(3-(trifluoromethyl)phenyl)pteridine-2,4,7-triamine
16		6-(m-tolyl)pteridine-2,4,7-triamine
17		6-(3-methoxyphenyl)pteridine-2,4,7-triamine
18		6-(3-aminophenyl)pteridine-2,4,7-triamine
19		6-(2-methoxyphenyl)pteridine-2,4,7-triamine
20		6-(2-fluorophenyl)pteridine-2,4,7-triamine
21		6-(2,6-dichlorophenyl)pteridine-2,4,7-triamine
22		6-(2,4-dichlorophenyl)pteridine-2,4,7-triamine
23		6-(3,4-dichlorophenyl)pteridine-2,4,7-triamine
24		6-(2,6-difluorophenyl)pteridine-2,4,7-triamine

No.	Compound structure	Compound name
25		6-(2,5-difluorophenyl)pteridine-2,4,7-triamine
26		6-(naphthalen-1-yl)pteridine-2,4,7-triamine
27		6-(pyridin-2-yl)pteridine-2,4,7-triamine
28		6-(benzo[b]thiophen-3-yl)pteridine-2,4,7-triamine
29		6-(thiophen-2-yl)pteridine-2,4,7-triamine
30		6-(pyridin-3-yl)pteridine-2,4,7-triamine
32		6,7-diphenylpteridine-2,4-diamine
33		6-phenylpteridine-2,4-diamine
34		3,3'-(2,4-diaminopteridine-6,7-diyl)diphenol
37		2-methyl-6-phenylpteridine-4,7-diamine
38		6-(3-fluorophenyl)-2-methylpteridine-4,7-diamine

No.	Compound structure	Compound name
39		2-(methylthio)-6-phenylpteridine-4,7-diamine
40		6-(3-fluorophenyl)-2-(methylthio)pteridine-4,7-diamine
50		6-(3-chlorophenyl)-4-ethoxypteridine-2,7-diamine
51		6-(3-chlorophenyl)-4-propoxypteridine-2,7-diamine
52		6-(3-chlorophenyl)-4-isopropoxypteridine-2,7-diamine
53		4-(allyloxy)-6-(3-chlorophenyl)pteridine-2,7-diamine
54		4-butoxy-6-(3-chlorophenyl)pteridine-2,7-diamine
55		4-(sec-butoxy)-6-(3-chlorophenyl)pteridine-2,7-diamine
56		6-(3-chlorophenyl)-4-(cyclopentyloxy)pteridine-2,7-diamine
57		6-(3-chlorophenyl)-4-(cyclopentylmethoxy)pteridine-2,7-diamine

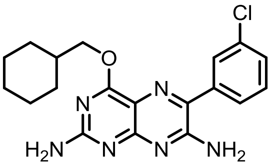
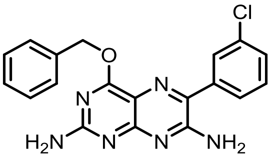
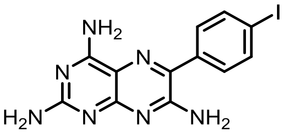
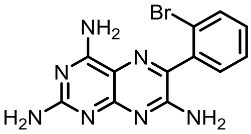
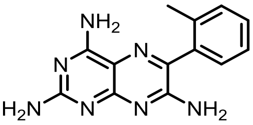
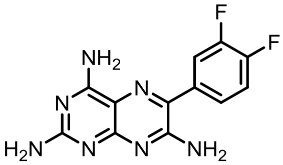
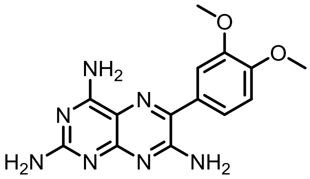
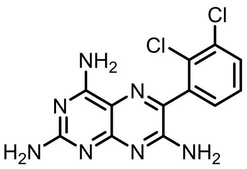
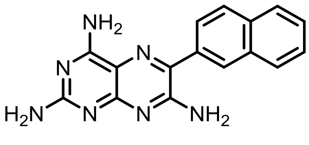
No.	Compound structure	Compound name
58		6-(3-chlorophenyl)-4-(cyclohexylmethoxy)pteridine-2,7-diamine
59		4-(benzyloxy)-6-(3-chlorophenyl)pteridine-2,7-diamine

Table A1. Chemical structures and IUPAC nomenclature of pteridine derivatives **2**-**59**.

No.	Compound structure	Compound name
60		6-(4-iodophenyl)pteridine-2,4,7-triamine
61		6-(2-bromophenyl)pteridine-2,4,7-triamine
62		6-(o-tolyl)pteridine-2,4,7-triamine
63		6-(3,4-difluorophenyl)pteridine-2,4,7-triamine
64		6-(3,4-dimethoxyphenyl)pteridine-2,4,7-triamine
65		6-(2,3-dichlorophenyl)pteridine-2,4,7-triamine
66		6-(naphthalen-2-yl)pteridine-2,4,7-triamine

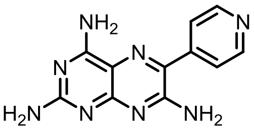
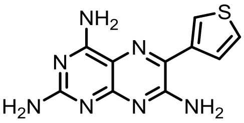
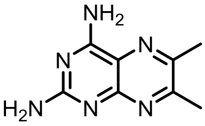
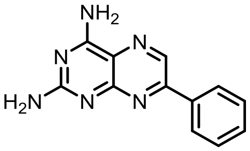
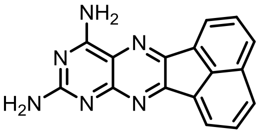
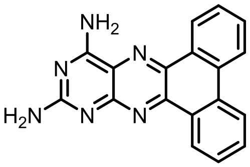
No.	Compound structure	Compound name
67		6-(pyridin-4-yl)pteridine-2,4,7-triamine
68		6-(thiophen-3-yl)pteridine-2,4,7-triamine
69		6,7-dimethylpteridine-2,4-diamine
70		7-phenylpteridine-2,4-diamine
71		acenaphtho[1,2-g]pteridine-9,11-diamine
72		phenanthro[9,10-g]pteridine-11,13-diamine

Table A2. Chemical structures and IUPAC nomenclature of pteridine derivatives **60**-**72**. These compounds were not investigated in biological context due to solubility issues.

3.2. Experimental part - Biology

3.2.1. Materials

Human basic fibroblast growth factor (bFGF) was purchased from Peprotech (Germany). Human leukemia inhibitory factor (LIF) was produced in-house.[148] KO-DMEM was acquired from Invitrogen. CHIR99021 and PD0325901 were purchased from Tocris Bioscience. Hoechst 33258 was obtained from Sigma. 4',6-diamidino-2-phenylindole (DAPI) was purchased from Invitrogen. Dulbecco’s Modified Eagle’s medium (DMEM), sodium pyruvate, penicillin / streptomycin, non-essential amino acids and trypsin were purchased from PAN Biotech (Germany). Fetal bovine serum was purchased from Invitrogen (Germany). Cell culture flasks were purchased from Sarstedt (Germany). Anti-CSNK1D (Abcam, mouse, ab85320), anti-CSNK1E (BD Biosciences, mouse, 610445), anti-CSNK1A1 (Santa Cruz, goat, C-19, sc-6477) and anti-REX1 (Abcam, ab28141) antibodies have been used according to the vendor’s recommended protocols. STK10 antibody (Santa Cruz, mouse, D-6, sc-398083) was purchased from Santa Cruz Biotechnology and used according to the vendors indications. 96-well plates were purchased from BD Biosciences (Germany). 384-well white plates with clear bottom were purchased from Corning (Germany). Protease and phosphatase inhibitor cocktails were obtained from Roche (Germany). LoBind Tubes (Eppendorf, Germany) were used for target engagement and proteomics experiments.

3.2.1.1. Cell lines

For testing reprogramming efficiencies of the pteridine-based derivatives the E3 EpiSC-GOF18 cell line was used which is derived at E5.5 from (C57BL/6 X DBA/2) X 129/Sv. In this case the Oct4 transgene corresponds to a 10 kbp fragment (genomic Oct4 fragment 18kbp: GOF18).[187] EpiSC medium (MEF-conditioned) contained Knockout(KO)-DMEM, 20 % Knockout - Serum Replacement (KOSR), 1 x L-glutamine, penicillin / streptomycin, non-essential amino acids, 100 μM 2-mercaptoethanol and 5 ng / mL bFGF (before and after conditioning over MEFs). Cells were cultured on fetal calf serum (FCS)-coated dishes and passaged with Accutase® (Life Technologies). mESCs or reverted EpiSCs were grown on MEFs using mESC medium containing Knockout (KO)-DMEM, 20 % Knockout - Serum Replacement, 1 x L-glutamine with penicillin / streptomycin, non-essential amino acids, 100 μM 2-mercaptoethanol and 2000 U / mL LIF. HS-578T (DSMZ-No. ACC-781), HT1080 (DSMZ-No. ACC-315), U2OS (DSMZ-No. ACC-785), SKBR3 (DSMZ-No. ACC-736), HCT116 (DSMZ-No. ACC-581), HEK293 (DSMZ-No. ACC-305), MCF7 (DSMZ-No. ACC-115) and HeLa (DSMZ-No. ACC-57) were purchased from Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures (Germany). ACHN (ATCC® CRL-1611™), MCF10A (ATCC® CRL-10317™), HEK293T (ATCC® CRL-3216™) and BT-549 (ATCC® HTB-122™) were obtained from ATCC (USA). HCT116 *TP53*$^{+/-}$ and HCT116 *TP53*$^{-/-}$ were a kind gift from Bert Vogelstein (Baltimore, USA). All cell lines were cultured following the growth media indicated by the vendors at 37°C in a 5 % CO_2 humidified atmosphere.

3.2.2. Methods

3.2.2.1. Reprogramming assay using mouse epiblast stem cells

(This experiment was fully performed by Dr. Damir J. Illich, Max Planck Institute for Molecular Biomedicine, Münster, Germany)

100,000 E3 EpiSC GOF18 cells were plated into 6 cm low attachment dishes on day 0. The reprogramming assay was performed in EpiSC medium CF1 MEF-conditioned KO-DMEM containing 20 % serum replacement and without additional bFGF. The compounds were added on day 0 at a concentration of 10 μM and the cells were incubated for 8 days at 37°C. Medium containing the compounds was changed on day 4 and 6. The basic EpiSC conversion medium was used as negative control, whereas a mixture of 3 μM CHIR99021 (GSK3β inhibitor), 1 μM PD0325901 (MEK1 inhibitor) and 2000 U / mL LIF (commonly referred to as 2i / LIF) was used as a positive control for reprogramming E3 EpiSC GOF18 cells. On day 8, cells were dissociated using trypsin. Approximately one million cells were resuspended in 500 μL of cell culture medium and the suspension was passed through a 40 μm cell strainer to remove debris and undissociated tissue. 4',6-diamidino-2-phenylindole (DAPI) was added at this stage to stain the nuclei of reprogrammed cells. Cells were subjected to fluorescence-activated cell sorting using a FACSAria cell sorter (BD Biosciences) to detect the selective expression of the Oct4-GFP construct in embryonic stem-like cells, i.e. reprogrammed cells. On day 8 of reprogramming, reverted E3 EpiSCs displayed reactivation of the Oct4-GFP transgene. The conversion efficiency was related to the initial found hit Triamterene (**TR**), whose efficiency was set to 1.

3.2.2.2. Immunocytochemistry

(This experiment was fully performed by Dr. Miao Zhang, Max Planck Institute for Molecular Biomedicine, Münster, Germany)

Cells seeded on coverslips were fixed with 4 % paraformaldehyde (PFA) for 10 min and then rinsed with 0.1 % Tween® 20 in phosphate-buffered saline (PBS) (PBS-T). Cells were permeabilized for 10 min with 0.5 % Triton X-100 in PBS at room temperature prior to incubation in a solution of 1 % bovine serum albumin (BSA), 5 % fetal calf serum and 2 % glycine in PBS-T for 1 h to block non-specific binding. Cells were then incubated with REX1 antibody (Abcam, ab28141, 1 μg / mL) overnight at 4°C. Afterwards, cells were washed in PBS-T and the secondary antibody (Alexa Fluor® 568, goat anti-rabbit IgG (H+L), 1:1000 dilution) was added for 1 h at room temperature. Cells were again washed in PBS-T and nuclei were stained with Hoechst 33342 for 5 min. Cells were then mounted onto glass slides. Images were acquired using a Leica fluorescence microscope.

3.2.2.3. Preparation of cell lysate for detecting CSNK1A1/D/E

HeLa, HCT116, MCF7, ACHN and HEK293T cells were cultured up to 80-85 % confluency. Growth media was removed, and cells were washed and collected by scrapping in ice-cold PBS. Cells were centrifuged for 5 min at 400 x g and lysed for 10 min on ice in a lysis buffer containing 50 mM PIPES (pH 7.4), 50 mM NaCl, 5 mM $MgCl_2$, 5 mM EGTA, 0.1 % NP40, 0.1 % Triton X-100, 0.1 % Tween® 20 and freshly added 0.1 % 2-mercaptethanol EDTA-free protease inhibitor cocktail (Complete EDTA free, Roche, Germany) and phosphatase inhibitors (PhosphoSTOP, Roche, Germany). Lysates were homogenized by passing them five times through 20G and 24G needles, respectively. Cellular debris was separated by centrifugation at 12000 x g for 20 min at 4°C. Clear cell lysates were frozen at -80°C. Protein concentration was determined using the Bradford reagent (Bio-Rad, Germany).

3.2.2.4. Immunoblotting

Cells were collected by scrapping and lysed following the procedure described above. Protein concentration determined using the Bradford reagent (Bio-Rad, Germany). Samples were adjusted to the same protein concentration prior to denaturation in Laemmli buffer at 95 °C for 8 min. Proteins were then separated by means of sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE) and transferred by means of semi-dry electroblotting to a polyvinylidene difluoride (PVDF) membrane (Thermo Scientific, Germany). Membranes were blocked using 5 % non-fat dry milk in Tris-buffered solution (10 mM Tris-HCl (pH 7.5), 100 mM NaCl, 0.1 % Tween® 20 for CSNK1E and 20 mM Tris-HCl (pH 7.6), 140 mM NaCl, 0.05 % Tween® 20 for CSNK1A1, CSNK1D and STK10) and were then incubated with the appropriate primary antibodies, diluted in blocking solution (CSNK1A1 1:300, CSNK1D 1:5000, CSNK1E 1:2000 and STK10 1:200) overnight at 4°C. Anti-α-tubulin (for CSNK1A1) or anti-GAPDH (for CSNK1D/E) antibodies were employed to detect α-tubulin or GAPDH as loading controls. Detection of bound primary antibodies was performed by incubating with appropriate secondary antibodies coupled to horseradish peroxidase (HRP, Pierce, USA) and SuperSignal West Pico or Femto Maximum Sensitivity Substrate (Pierce, USA). For reprobing, membranes were stripped in Restore Western Blot Stripping Buffer (ThermoScientific, USA) for 10 min at room temperature according to the vendor's instructions.

3.2.2.5. Plasmid construction, protein expression and purification

(The plasmid construction, protein expression and purification was fully performed at the Dortmund Protein Facility, Max Planck Institute of Molecular Physiology, Dortmund, Germany)

Fully sequenced human cDNA constructs for CSNK1D (former cat. no. MHS4771-202829011; replaced by MHS6278-202829011, clone ID 3546633, accession no. BC003558) and CSNK1E (cat. no. MHS6278-202830217, clone ID 3937782, accession no. BC006490) were purchased from Open Biosystems - Dharmacon (GE Healthcare). The nucleotide sequence corresponding to amino acids 1-294 of the human CSNK1D and CSNK1E cDNA was amplified using the oligonucleotides:
5'-GAAGTTCTGTTTCAGGGTCCCATGGAGCTGAGAGTCGGGAACA-3'
5'-CTTTGTTAGCAGCCGGATCATTATTTGAGCATGTTCCAGTCGAACACG-3'
for CSNK1D and
5'-GCGAACAGATCGGTGGTATGGAGCTACGTGTGGGGAAC-3'
5'-ATTTCAGCATGTTCCAGTCAAAGACGTA-3'
for CSNK1E and subsequently subcloned in pET19-nHis-3C for CSNK1D and pOPIN-His-Sumo for CSNK1E. pET19-nHis-3C-CSNK1D_1-294 and pOPIN-His-Sumo-CSNK1E_1-294 were transformed in the *E. coli* strain BL21 DE3 RIL(K+). Bacteria were grown in SB medium (3.2 % peptone, 2 % yeast extract and 0.5 % NaCl) at 37°C. Expression was induced with 400 μM IPTG at OD_{600} of 0.6. Cells were grown at 25°C for 18 h. Harvested cells were suspended in lysis buffer (50 mM Tris, pH 8.0, 400 mM NaCl, 5 % glycerol, 0.1 % Triton X-100, 1 mM DTT, 5 mM imidazol and protease inhibitors) and lysis was performed by sonication. After centrifugation the supernatant was loaded on a Ni-NTA column (Qiagen). After washing with 50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 10 % glycerol and 5 mM imidazol, proteins were eluted with 50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 10 % glycerol and 30 mM imidazol. In collected fractions proteins were digested using Prescission protease

/ Sumo protease and were further purified by size exclusion chromatography using a Sephadex G200 column with 25 mM Hepes pH 7.0, 150 mM NaCl, 2 mM TCEP, 10 mM $MgCl_2$ and 5 % glycerol for CSNK1D and 25 mM Hepes, pH 7.5, 150 mM NaCl, 1 mM TCEP, 10 mM $MgCl_2$ and 5 % glycerol for CSNK1E. For both proteins, the buffer was exchanged to CK1 dilution buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1 mM EDTA, 0.25 mM DTT, 25 % glycerol). To this end, Zeba Spin Desalting Column (7K MWCO, 2 mL; Thermo Scientific) was centrifuged at 1000 x g for 2 min at room temperature. The column was equilibrated three times with 1.2 mL ice-cold CK1 dilution buffer at 1000 x g for 2 min at room temperature. After the final centrifugation step, the kinase containing fraction was added and centrifuged for 2 min at 1000 x g and room temperature. The newly obtained kinase containing solution was collected and protein concentration was determined using the Bradford reagent (Bio-Rad Protein Assay Kit I, Bio-Rad, cat no 5000001) and adjusted to approx. 1 μg / μL. Proteins were snap-frozen in liquid nitrogen and stored at -80°C.

3.2.2.6. Crystal structure determinations

(This experiment was performed by Arthur Porfetye, Max Planck Institute of Molecular Physiology, Dortmund, Germany)

Crystals of CSNK1D were obtained at 4°C in hanging drops containing 0.1 M Li_2SO_4, 0.7 - 0.8 M Na-K tartrate and 0.1 M CHES (pH 9.5). Protein concentration used was 7.45 mg / mL in CK1 dilution buffer. Crystals appeared after 21 days as small ellipsoids and were mounted in a thin film of mother liquor without additional cryo protectant. Structures of CSNK1D in complex with Epiblastin A and derivative **15** were obtained by soaking crystals overnight in mother liquor containing 4 mM of Epiblastin A and derivative **15** (40 % DMSO concentration in the final drop). The crystals belonged to the space group 152 ($P3_121$) with a unit cell parameters of a = b = 64.5 Å, c = 151.8 Å, $\alpha = \beta = 90^\circ$, $\gamma = 120^\circ$ and one molecule in the asymmetric unit. Diffraction data was collected at PXII at Swiss Light Source (SLS) at Paul-Scherrer-Institute, Villigen, Switzerland using a Pilatus 6M detector. The data was processed and scaled using XDS / XSCALE[188] and solved by molecular replacement with PHASER[189] from the CCP4-suite[190] using PDB entry code 3UYS as template. Model building and refinement was carried out using REFMAC5[191] and coot.[192] The data collection and refinement parameters of the apo-CSNK1D and CSNK1D in complex with Epiblastin A and derivative **15** crystal structures can be found in **Table A3**.

	Apo-CSNK1D	CSNK1D - Epiblastin A complex	CSNK1D - derivative **15** complex	CSNK1D - derivative **15** complex data set 2
PDB ID				
Data collection				
Wavelength (Å)	0.91639	0.92019	1.76995	0.91883
Resolution (Å)	55.89 - 1.91 (1.95 -1.91)	50 - 2.59 (2.66 - 2.59)	45.74 - 2.76 (2.83 - 2.76)	2.8Å (2.87-2.8)
Space group	$P3_1 2 1$	$P3_1 2 1$	$P3_1 2 1$	$P3_1 2 1$
Unit cell				
a, b, c (Å)	64.5, 64.5, 151.8	65.1, 65.1, 151.6	65.4, 65,4 151.7	65.4, 65.4, 151.7
α, β, γ (˚)	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 120
Total reflections	542,287 (29,503)	223,430 (16,573)	104,020 (15,080)	189,434 (13,879)
Unique reflections	53,992 (3,340)	22,311 (1,668)	18,722 (2,960)	17,797 (1,303)
R_{merge} [a]	0.115 (1.323)	0.148 (1.546)	0.087 (0.897)	0.151 (1.2)
Redundancy	10.0	10.0	5.6	10.6 (10.7)
Completeness (%)	96.0 (83.9)	99.9 (99.7)	99.6 (95.7)	99.5 (99.3)
I/σI	13.45 (1.64)	12.23 (1.59)	11.73 (1.74)	10.64 (1.84)
$CC_{1/2}$	99.8 (87.1)	99.9 (86.7)	99.8 (91.6)	99.9 (89.2)
Refinement				
Resolution (Å)	55.89-1.91 (1.96 -1.91)	56.38-2.59 (2.66-2.59)	45.47-2.76 (2.83-2.76)	56.72-2.8 (2.87-2.8)
No. reflections used	27,094	11,526	9,718	9,225
R_{work} (%)	20.2 (53.4)	23.2 (41.1)	20.9 (37.9)	20.6 (31.3)
R_{free} (%)	23.6 (53.2)	29.0 (43.3)	30.1 (45.5)	31.6 (45.7)
No. atoms / Mean B-values (Å^2)				
Protein	2373 / 41	2320 / 74	2307 / 92	2243 / 90
Inhibitor / Solute	31 / 75	41 / 98	41 / 82	41 / 107
Water	243 / 48	15 / 59	34 / 105	33 / 78
Rms deviation				
Bond length (Å)	0.007	0.012	0.01	0.009
Bond Angle (˚)	1.177	1.318	1.27	1.303
Ramachandran (%)				
favoured	97.92	96.43	98.92	98.89
accepted	2.08	3.57	1.08	1.11
outlier	0.0	0.0	0.0	0.0

Table A3. Crystallographic data collection and refinement statistics for the crystal structures of apo-CSNK1D, and CSNK1D in complex with Epiblastin A and derivative **15**, respectively. Data set 2 refers to the detection of the anomalous scattering of the bromine atom from derivative **15**.

3.2.2.7. Cell lysate preparation for target engagement and chemical proteomics

HCT116 cells were grown up to 80-90 % confluency. Cells were washed and collected by scraping using ice-cold PBS. Cells from multiple T175 flasks were combined and centrifuged for 5 min at 500 x g and 4°C. Cell pellets were either directly frozen (and stored at -80°C) or subsequently lysed in ice-cold Pierce IP Lysis buffer (25 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 mM EDTA, 1 % NP-40 and 5 % glycerol) supplemented with Halt Protease and Phosphatase Inhibitor Cocktail (Thermo Scientific, prod. no. 78440) for 10 min on ice. To assure complete cell lysis samples were gently vortexed every three min until a homogeneous mixture was obtained. The resulting homogenate was centrifuged for 5 min at 16000 x g and 4°C and supernatant were carefully transferred into a new vial. For the buffer exchange procedure, the protocol suggested by the vendor was applied. Zeba Spin Desalting Columns (7K MWCO, 5 mL; Thermo Scientific, prod. no. 89891) were centrifuged at 1000 x g for 2 min at room temperature. The column was equilibrated with ice-cold Kinase buffer (20 mM Hepes pH 7.4, 150 mM NaCl, 0,1% Triton X-100) and centrifuged at 1000 x g for 2 min at room temperature prior to addition of the cell lysate and centrifugation for 2 min at 1000 x g and room temperature. The buffer-exchanged lysate was and the protein concentration was adjusted to approximately 2 mg / mL with ice-cold Kinase buffer. Lysates were subsequently stored at -80°C.

3.2.2.8. Kinase profiling at Eurofins (former Milipore)

Single-point measurement analyses of Epiblastin A against a panel of 123 kinases was performed at Eurofins (former Milipore) using a radiometric-based assay and K_m-ATP (K_m is defined as the concentration of ATP that allows half-maximal reaction velocity) concentration for each individual kinase. Dose-dependent investigations of Epiblastin A against BRSK1, EGFR, EEF2K, MKNK2, RIPK2 and STK10 have been performed under the same experimental setup, i.e. at K_m-ATP value. Half-maximal inhibitory concentrations (IC_{50}) were obtained by fitting the obtained data with three parameter nonlinear regression analysis using GraphPad Prism version 6.00 for Windows (GraphPad Software, USA). For more details related to the experimental conditions employed at Eurofins please consult the following website:

http://www.eurofins.com/pharma-services/pharma-discovery-services/services/in-vitro-pharmacology/kinases/biochemical.aspx

3.2.2.9. Determination of CSNK1D/E activity using the KinaseGlo® Plus luminescent assay

(The assay conditions were preliminary optimized by Dr. Yasushi Takemoto, RIKEN-Max Planck Joint Research Center for Systems Chemical Biology, Wako, Japan)

A stock solution of 5 % DMSO (BioReagent for molecular biology, Sigma Aldrich) in water was prepared. The CSNK1 substrate peptide RRKDLHDDEEDEAMSITA (Jena Bioscience, cat. no. PE-206, 2 mM stock solution in water) was diluted in water to a final concentration of 250 μM. CSNK1D or CSNK1E kinases were diluted with CK1 dilution buffer to 50 ng / μL. In addition, fresh Reaction buffer was prepared containing 250 mM Tris-HCl (pH 7.5), 50 mM $MgCl_2$, 5 mM DTT and 0.5 mg / mL bovine serum albumin (BSA). Compounds were serially diluted in 5 % DMSO solution to the desired concentrations (maximal concentration employed in the assay was 50 μM). Additional samples included blank sample containing Reaction buffer, water and CSNK1D / CSNK1E kinase and control sample containing Reaction buffer, CSNK1 substrate peptide and CSNK1D / CSNK1E kinase. CSNK1 substrate peptide, Reaction buffer and CSNK1D / CSNK1E kinase stock solutions were mixed in 1:1:1 ratio. 6 μL of this mixture was added to 384-well plates (384 well low volume white round bottom polystyrene NBS™, Corning®, prod. no. 3673). 2 μL of the test compound stock solutions were subsequently added to the wells. The same volume of the 5 % DMSO solution was added to the blank and control samples, respectively (final DMSO concentration in the assay was 1 %). The content of the plate was centrifuged for 1 min at 750 x g and room temperature. The plate was incubated under rotation / shaking at room temperature for 5 min, followed by 25 min incubation at room temperature. The kinase reaction was initiated by adding 2 μL ATP (Ultra Pure ATP, Promega, product no V703A-C, 100 mM stock solution in water) from a 50 μM stock solution in water. The content of the plate was centrifuged for 1 min at 750 rpm and room temperature, followed by shaking for 5 min at room temperature and incubation at room temperature for 55 min. Afterwards, 10 μL of the Kinase-Glo® Plus reagent (Promega, cat. no. V3771) was added to each well and the plate was incubated at room temperature for 5 min. Luminescence was recorded on an Infinite® M200 plate reader (Tecan, Austria) (integration time 1000 ms, settle time 5 ms). CSNK1D / CSNK1E kinase activity (% of control) was calculated as follows:

Kinase activity (%) = ((<luminescence (blank sample)> - <luminescence (compound sample)>) / (<luminescence (blank sample)> - <luminescence (control sample)>)) x 100.

Half-maximal inhibitory concentrations (IC_{50}) were generated by fitting the obtained data with three parameters nonlinear regression analysis using GraphPad Prism version 6.00 for Windows (GraphPad Software, USA).

3.2.2.10. CSNK1A profiling at Reaction Biology

Single-point measurements analyses and subsequent dose-dependent investigations of selected pteridine derivatives for inhibition of CSNK1A1 were performed at Reaction Biology (USA) using a radiometric-based assay and ATP concentration of 10 μM. For more details related to the experimental conditions employed please consult the following websites: http://www.reactionbiology.com/webapps/site/KinaseDetail.aspx
http://www.reactionbiology.com/webapps/site/Kinase_Assay_Protocol.aspx
Half-maximal inhibitory concentrations (IC_{50}) were generated by fitting the obtained data with three parameter nonlinear regression analysis using GraphPad Prism version 6.00 for Windows (GraphPad Software, USA).

3.2.2.11. Detection of target engagement using the ActivX^{TM} ATP probe

HCT116 cell lysate (approx. 1 mg total protein) was thawed on ice prior to activation by adding 10 μL of 1 M $MgCl_2$ stock solution to each sample. Samples were vortexed gently and incubated for 1 min at room temperature. 1 % of the entire volume was removed and replaced by the same volume of DMSO (BioReagent for molecular biology, Sigma Aldrich) or Epiblastin A stock solution in 100 % DMSO, respectively (0.1, 1 and 10 mM Epiblastin A initial stock solutions were used to yield 1, 10 and 100 μM Epiblastin A final concentration in the cell lysate; DMSO content in all samples was 1 %). Samples were gently vortexed and incubated at room temperature for 10 minutes. Two vials containing 12.6 μg of the ActivX^{TM} ATP probe each (Life Technologies, cat. no. 88311; stored at -80°C) were pre-equilibrated at room temperature. 22.5 μL water was added to one vial and the content from the first vial was transferred to the second vial (final solution contains 2 x 12.6 μg ActivX^{TM} ATP probe dissolved in 22.5 μL water). 5 μL from the final solution was transferred to the DMSO and Epiblastin A-treated samples, respectively. The content was gently mixed and incubated at room temperature for 10 min. The solution was diluted 1:2 in 8 M urea stock solution in Pierce IP lysis buffer (25 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 mM EDTA, 1 % NP-40 and 5 % glycerol), followed by the addition of 45 μL of high capacity streptavidin agarose resin

(50 % slurry; Pierce™ Streptavidin Agarose**,** cat no 20347). The resulting mixture was incubated for 1 hour at room temperature with overhead rotation (speed 10; Rotator SB3, Stuart). Samples were centrifuged at 1000 x g and room temperature for 1 min. The supernatant was carefully replaced by the same volume of 4 M urea in Pierce IP lysis buffer (same volume that has been added in the case of 8 M urea dilution; see previous step). This mixture was transferred into a spin column and centrifuged at 1000 x g and room temperature for 1 min. The washing step with 4 M urea in Pierce IP lysis buffer was repeated twice. Finally, 45 μL of a 2 x Laemmli buffer (100 mM Tris pH = 6.8, 7 % SDS, 200 mM DTT, 0.014 % Bromphenol blue) was added to the separated beads on the spin column. The spin column was closed, placed in a new vial and proteins were denatured for 10 min at 95ºC under shaking. The vials were left to cool down at room temperature and the content of the spin column was collected in a new vial at 5000 x g for 3 min at room temperature. Target engagement experiments were carried out by varying compound incubation time and / or the incubation time for the ActivX^{TM} ATP probes, respectively with respect to each kinase. Samples were then analyzed by means of immunoblotting or mass spectrometry.

3.2.2.12. Chemical proteomics experiment based on the ActivX^{TM} ATP probe

Same protocol was applied as previously described for the detection of target engagement. In addition, a BSA control was prepared by adding 1 μL of 5 pmol / μL BSA stock solution in a separate vial together with 45 μL of the 2 x Laemmli buffer. A total of 7 samples were boiled for 10 min at 95ºC with shaking (500 rpm) (Thermomixer® comfort, Eppendorf). Samples were left to cool down at room temperature and centrifuged in new vials at 5000 x g for 3 min at room temperature. Samples were loaded onto Mini-Protean TGX precast gels (4-20 %). 2 x Laemmli buffer was loaded in a separate well as an additional control for the proteomics procedure. Proteins were separated for 20 min at 80 V. The gel washed overnight with a mixture containing water, EtOH (ethanol) and acetic acid in a 5:4:1 ratio. The gel was cut in pieces with a scalpel and transferred into new LoBind tubes. Gel pieces were washed with 200 μL washing solution 1 (WS1, 25 mM ammonium bicarbonate : acetonitrile, 3:1), tubes were briefly centrifuged and incubated for 30 min at 37°C with shaking at 600 rpm. Supernatants were removed and the procedure was repeated by adding 200 μL washing solution 2 (WS2, 25 mM ammonium bicarbonate : acetonitrile; 1:1) under the same conditions and incubation for 15 min. The solution was removed and 200 μL reducing solution was added (50 mM dithiothreitol in 25 mM ammonium bicarbonate, always freshly prepared). Vials were briefly centrifuged and incubated for 45 min at 37°C. The reducing solution was then exchanged for 200 µL alkylating solution (55 mM iodacetamide in 25 mM ammonium bicarbonate). Samples were briefly centrifuged and the content was incubated under shaking for 1 hour at room temperature in the dark. The alkylating solution was removed and gel pieces were washed twice for 15 minutes with WS2 at room temperature. Afterwards, the solutions were removed as completely as possible. The gel pieces were dehydrated by adding 100 µL acetonitrile for 10 minutes at room temperature. At this stage the gel turned white. The acetonitrile was removed and the gel pieces were left to dry for 10 min. Subsequently, 50 µL digest solution was added (0.01 µg / µl Trypsin in 25 mM ammonium bicarbonate solution). Vials were briefly centrifuged and incubated for 15 minutes at room temperature. 75 µL 25 mM ammonium bicarbonate was subsequently added and vials were again briefly centrifuged and incubated overnight at 30°C with shaking (350 rpm). 10 µL 10 % trifluoroacetic acid was added to the gel pieces and vials were briefly centrifuged and incubated for 30 min in a sonicator bath on ice. 75 µL acetonitrile was added to the gel pieces, vials were briefly centrifuged and were left to shake for 15 minutes at 350 rpm at room temperature. The supernatant was transferred into new vials and 75 µL acetonitrile was added to the gel pieces, the vials were briefly centrifuged and subsequently incubated for additional 15 minutes at room temperature with shaking (350 rpm). Both acetonitrile containing fractions were combined and vacuum concentrated. Samples were stored at -20°C.

3.2.2.13. Mass spectrometric analysis

For protein identification tryptic peptides were separated and analyzed by nano-HPLC/MS/MS. The separations were carried out on UltiMateTM 3000 RSLCnano system (Dionex, Germany). The MS and MS / MS experiments were carried out on a Q Exactive™ Hybrid Quadrupole-Orbitrap Mass Spectrometer equipped with a nano-spray source (Nanospray Flex Ion Source, Thermo Scientific). All solvents were LC-MS grade. The lyophilized tryptic peptides were dissolved in 20 µl 0.1 % TFA in water. 3 µL of sample were injected onto a pre-column cartridge (5 µm, 100 Å, 300 µm ID * 5 mm, Dionex, Germany) using 0.1 % TFA in water as eluent with a flow rate of 30 µL/min. Desalting was performed for 5 min with eluent flow to waste followed by back-flushing of the sample during the whole analysis from the pre-column to the PepMap100 RSLC C18 nano-HPLC column (2 µm, 100

Å, 75 µm ID × 25 cm, nanoViper, Dionex, Germany) using a linear gradient starting with 95 % water containing 0.1 % formic acid / 5 % acetonitrile containing 0.1 % formic acid and increasing to 70 % water containing 0.1 % formic acid / 30 % acetonitrile containing 0.1 % formic acid after 95 min using a flow rate of 300 nL / min.

The nano-HPLC was online coupled to the Quadrupole-Orbitrap Mass Spectrometer using a standard coated SilicaTip (ID 20 µm, Tip-ID 10 µM, New Objective, Woburn, MA, USA). Mass range of m/z 300 to 1650 was acquired with a resolution of 70000 for full scan, followed by up to ten high energy collision dissociation (HCD) MS / MS scans of the most intense at least doubly charged ions.

Data evaluation was performed using MaxQuant software[193] (v.1.4.1.2) including the Andromeda search algorithm and searching the human reference proteome of the Uniprot database. The search was performed for full enzymatic trypsin cleavages allowing two miscleavages. For protein modifications carbamidomethylation was chosen as fixed and oxidation of methionine and acetylation of the N-terminus as variable modifications. The mass accuracy for full mass spectra was set to 5 ppm and for MS / MS spectra 20 ppm. The false discovery rates for peptide and protein identification were set to 1 %. Only proteins for which at least two peptides were quantified were chosen for further validation. Relative quantification of proteins was carried out using the label-free quantification algorithm implemented in MaxQuant. Briefly, samples resulting from affinity enrichments with the active molecule bound to the solid support were grouped together and those resulting from similar enrichment using control molecules as well. All experiments were performed in technical triplicates. Label-free quantification (LFQ) intensities were logarithmized (log2) and proteins which were not three times quantified in at least one of the groups were filtered off. Missing values were imputed using small normal distributed values and a median test was performed. Proteins which were statistically significant outliers were considered as hits.

3.2.2.14. WST-1 cell proliferation assays

The WST-1 cell proliferation assay (Roche, cat no 11644807001) was performed in 96-well plates (clear and sterile; BD Biosciences, Germany) as follows. HCT116 cells (80-85 % confluent) were seeded in quadruplicate (4000 cells / well, final volume 100 µL) and incubated overnight at 37°C. Afterwards, the culture medium was replaced with fresh growth medium previously supplemented with DMSO and various concentrations of Epiblastin A, respectively. Cells were then incubated for 24 h at 37°C prior to addition of the WST-1 reagent following the indications from the vendor. Samples were incubated at 37°C between 30 min and 60 min. To detect the conversion of WST-1 to formazan, absorbance at 450 nm (and reference wavelength of 620 nm) was recorded on an Infinite® M200 plate reader (Tecan, Austria). 12 different Epiblastin A concentrations were employed for the final determination of the half-maximal growth inhibitory (GI_{50}) concentration. The absorbance corresponding to the background control (containing culture media without cells and WST-1 reagent) recorded at 620 nm was subtracted from the absorbance values determined at 450 nm for each compound-treated sample including DMSO. The percentage of viable cells was calculated by relating absorbance of the compound-treated samples to the absorbance obtained for cells treated with DMSO, which was set to 100 % cell viability. Half-maximal growth inhibitory concentrations values were determined by fitting the generated data with the four parameters equation using GraphPad Prism version 6.00 for Windows (GraphPad Software, USA). Analogous procedure was followed for analyzing the cell viability of other cell lines upon treatment with Epiblastin A or D4476. The number of cells plated for 24 h incubation time was as follows: U2OS: 2000 cells / well, BT549: 300 cells / well, HS578T: 1500 cells / well, HT1080: 500 cells / well, MCF10A: 1500 cells / well, SKBR3: 4000 cells / well. HCT116 cells that lack one or both *TP53* alleles (HCT116 $TP53^{-/+}$ and HCT116 $TP53^{-/-}$) were plated as follows: 4000 cells / well for 24 h incubation and 3000 cells / well for 48 h incubation.

3.2.2.15. CellTiter-Glo® viability assay

This assay follows an analogous experimental procedure as previously described for WST-1 cell proliferation assay. After overnight incubation with the compound under investigation, 100 μL CellTiter-Glo® reagent (prod. no. G7570, Promega) was added. The luminescent signal generated was recorded on an Infinite® M200 plate reader (Tecan, Austria). The percentage of viable cells was calculated by relating the luminescence of the compound-treated samples to the luminescent signal generated by the DMSO treated sample, which was set to 100 % cell viability. Half-maximal growth inhibitory concentrations (GI_{50}) values were determined by fitting the generated data with the four parameters equation using GraphPad Prism version 6.00 for Windows (GraphPad Software, USA).For more details regarding the CellTiter-Glo® cell viability assay please consult the following website: https://www.promega.de/en/products/cell-health-and-metabolism/cell-viability-assays/celltiter_glo-luminescent-cell-viability-assay/

3.2.2.16. CellTox™ Green cytotoxicity assay

This assay follows an analogous experimental procedure as previously described for WST-1 and CellTiter-Glo® assays. After overnight incubation with the compound under investigation, 100 µL CellTox™ Green reagent (prod. no. G8741, Promega) was added following the Endpoint method (2 x) described by the vendor. The fluorescence emission was recorded at 515 nm upon excitation with 490 nm light using an Infinite® M200 plate reader (Tecan, Austria). The percentage of viable cells was calculated by relating the signal generated by the compound-treated sample to the emission detected from the cells treated with DMSO, which was set to 100 % cell viability. Half-maximal growth inhibitory concentrations (GI_{50}) values were determined by fitting the generated data with the four parameters equation using GraphPad Prism version 6.00 for Windows (GraphPad Software, USA).For more details regarding the CellTox™ Green cell viability assay please consult the following website:
https://www.promega.de/en/products/cell-health-and-metabolism/cytotoxicity-assays/real-time-cytotoxicity-assay/celltox-green-cytotoxicity-assay/

3.2.2.17. Determination of the cell viability using Tryphan blue

The assay was performed in 6-well plates (clear and sterile; prod. no. 83.3920, Sarstedt, Germany) as follows. HCT116 cells (80-85 % confluent) were seeded in duplicate (35,000 cells / well for 24 h experiment and 25,000 cells / well for 48 h; final volume 2 mL) and seeded overnight at 37°C. Afterwards, the media was replaced with fresh growth medium supplemented with DMSO and various concentrations of Epiblastin A. Cells were then incubated for 24 h and 48 h, respectively at 37°C. Afterwards, the media was removed, cells were washed twice with 1 mL PBS, tripsinized and collected with 500 µL of fresh growth medium. 10 µL of this solution was combined with 10 µL of 0.4 % Trypan blue stain (prod. no. T10282, Invitrogen) and incubated at room temperature for 2 min. 10 µL of the final solution was added to counting chamber slide (prod. no. C10228, Thermo Scientific) and the number of viable cells was assessed using Countess™ Automated Cell Counter (Thermo Scientific, Germany). The percentage of viable cells upon compound treatment was calculated by relating to the number of cells determined for the DMSO treated sample, which was set to 100 % cell viability.

4. Results

4.1. Structure activity relationship (SAR) studies

To avoid any biases associated with either the origin of the mouse epiblast stem cell line or the population heterogeneity previously reported (see section *1.1.10. Reprogramming mEpiSCs to mES-like cells – state of the art*), small molecules capable of reprogramming late-stage mEpiSCs are required. Moreover, these chemical agents would desirably act on refractive mEpiSC lines also, namely cell lines that do not undergo spontaneous conversion to ESC-like cells upon 2i / LIF media switching. The identification of such small molecules offers a much more generalized reprogramming method coupled with the potential to identify the signaling pathways responsible for the cellular transition mechanism.

Following the preliminary results previously mentioned (see section *1.1.11. Reprogramming late-stage mEpiSCs to mES-like cells – preliminary results*), pteridine derivatives with higher reprogramming efficiencies compared to the initial found hit **TR** were sought. In order to determine how the substitution pattern on the pteridine core impacts the conversion ability of the starting cell population, structural variations at different positions such as C2, C6 and C7, respectively were explored (see **Figure 8**). All derivatives were tested for reprogramming late-stage Oct4-GFP-negative mEpiSCs in combination with **PD**, both applied at 10 µM concentration for 7 days following the previously described conditions.[148,149] The assay was fully performed by Dr. Damir J. Illich (Max Planck Institute for Molecular Biomedicine, Münster, Germany).

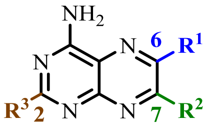
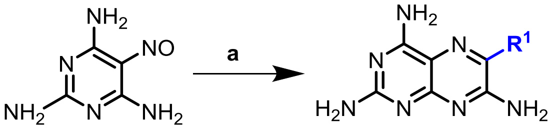


Figure 8. Derivatization positions on the pteridine scaffold employed for the synthesis of pteridine-based compound library.

First, the substitution pattern on the benzene ring attached to the pteridine scaffold at position C6 was explored because of the synthetic accessibility of analogues from readily available starting materials and the observation that seemingly small modifications produce marked perturbations of the activity of various enzymes.[185,194,195] Therefore, parameters such as electron density on the phenyl ring as well as the steric demand of the substituents might be crucial for modulation of enzymatic activity. Moreover, the dihedral angle at position C6 might also play an important role in fine-tuning potency and / or selectivity. For compound collection synthesis, commercially available 5-nitroso-2,4,6-triaminopyrimidine **1** was subjected to reaction with a broad range of phenylacetonitriles to yield derivatives **2** – **30** in a single step, using a procedure reported in the literature.[182,184,185,186] (see **Scheme 1**).



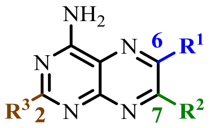
1 derivatives **2-30**

a. Phenylacetonitrile (1.1 eq), NaH (1.1 eq), 2-ethoxyethanol, reflux (1 - 3 h)

Scheme 1. Synthesis route to access pteridine-derived compounds bearing substitutions at position C6.

Even small modifications of the phenyl ring affected the conversion efficiency significantly as shown by entries **2-30** in **Table 2**. For example, a fluorine atom in the para-position as in derivative **3** (entry **3**) diminished the reprogramming activity to approximately half compared to **TR** (entry **2**; the relative biological activity of **TR** was arbitrarily set to 1). Introduction of bulkier substituents like methyl-, amino- or methoxy- in derivatives **6**, **7** and **5** respectively, led to a larger drop or complete loss of the reprogramming activity (compare entry **2** with entries **6**, **7** and **5**, respectively). Thus, it could be demonstrated that an unsubstituted para-position on the phenyl ring at position C6 is essential for maintaining the biological activity. In addition, varying the substitution pattern at the meta-position was well accepted. For example, meta-fluoro (entry **13**) and meta-methyl (entry **14**) containing derivatives were more active than **TR** (compare entry **2** with entries **13** and **14** respectively). The introduction of a chlorine atom in the same position yielded derivative **12** (entry **12**) that exhibited an 8-fold increased reprogramming efficiency compared to the guiding compound (compare entry **2** with entry **12**). A slightly larger halogen substituent such as bromine generated derivative **15** (entry **15**) which is equipotent to **TR**. Introduction of an alkoxy, trifluoromethyl or amino group as in derivatives **17**, **16** and **18** (entries **17**, **16** and **18** respectively) led to a reduction of the biological activity compared to **TR**. Derivatization of the ortho-position was also well tolerated, as seen in ortho-fluorine and ortho-methoxy substitution of compounds **20** and **19** (entries **20** and **19**) which showed activities comparable to the parent molecule **TR**.

Overall, the structure-activity relationship study revealed that meta- and ortho-substitutions were tolerated and in some cases even led to higher biological activity compared with the parent hit. However, even minor structural changes of the para-position, such as the introduction of a fluorine atom, nearly abolished the conversion efficiency.



Nr.	**R^1**	**R^2**	**R^3**	**Efficiency**
2	-H	$-NH_2$	$-NH_2$	1
3	p-F-Ph	$-NH_2$	$-NH_2$	0.6
4	p-OAllyl-Ph	$-NH_2$	$-NH_2$	0.1
5	p-OCH_3-Ph	$-NH_2$	$-NH_2$	0.1
6	p-CH_3-Ph	$-NH_2$	$-NH_2$	0.07
7	p-NH_2-Ph	$-NH_2$	$-NH_2$	0
8	p-NO_2-Ph	$-NH_2$	$-NH_2$	0
9	p-iPr-Ph	$-NH_2$	$-NH_2$	0
10	p-Ph-Ph	$-NH_2$	$-NH_2$	0
11	p-OBz-Ph	$-NH_2$	$-NH_2$	0
12	m-Cl--Ph	$-NH_2$	$-NH_2$	8.4
13	m-F-Ph	$-NH_2$	$-NH_2$	3.1
14	m-Me-Ph	$-NH_2$	$-NH_2$	1.6
15	m-Br-Ph	$-NH_2$	$-NH_2$	1.1
16	m-CF_3-Ph	$-NH_2$	$-NH_2$	0.7
17	m-OCH_3-Ph	$-NH_2$	$-NH_2$	0.3
18	m-NH_2-Ph	$-NH_2$	$-NH_2$	0.2
19	o-OCH_3-Ph	$-NH_2$	$-NH_2$	1.3
20	o-F-Ph	$-NH_2$	$-NH_2$	1.3
21	o,o‘-diCl-Ph	$-NH_2$	$-NH_2$	2.2
22	o,p-diCl-Ph	$-NH_2$	$-NH_2$	0.5
23	m,p-diCl-Ph	$-NH_2$	$-NH_2$	1.6
24	o,o‘-diF-Ph	$-NH_2$	$-NH_2$	2.3
25	m,o‘-diF-Ph	$-NH_2$	$-NH_2$	0.8
26	Naphtalen-1-yl	$-NH_2$	$-NH_2$	2.6
27	Pyridin-2-yl	$-NH_2$	$-NH_2$	0.5
28	Benzo[b] thiophen-3-yl	$-NH_2$	$-NH_2$	0.4
29	Thiophen-2-yl	$-NH_2$	$-NH_2$	0.1
30	Pyridin-3-yl	$-NH_2$	$-NH_2$	0.01

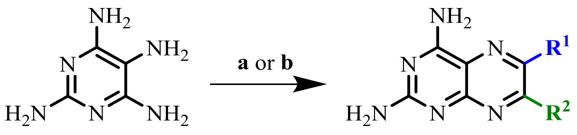
Table 2. Reprogramming efficiency of late-stage mEpiSCs upon treatment with derivatives **2-30** and **PD** at 10 µM concentration for 7 days. The relative reprogramming activity of **TR** (entry **2**) was arbitrary set to 1. **PD** = PD0325901 (MEK1 inhibitor). Data are representative of three independent experiments.

Given the strong influence of phenyl mono-derivatization on the biological activity, various disubstituted derivatives were subsequently generated following the same synthetic procedure. Not surprisingly, the disubstitution pattern on the phenyl ring had a strong impact on the reprogramming yield. For example, two ortho-fluorine or ortho-chlorine atoms as shown for derivatives **24** and **21** (entries **24** and **21**) are beneficial as compared to **TR** (compare entry **2** with entries **24** and **21**, respectively). However, introduction of two chlorines in para- and meta-position as in compound **23** (entry **23**) only slightly increased the reprogramming efficiency. This is in agreement with the previous observation made for the monosubstituted analogs, i.e. the positive influence of the meta-chloro derivatization appears to be negatively balanced by the para-substituent. The remaining derivatives **22** and **25** (entries **22** and **25**) failed to generate more active compounds than the initial found hit (compare entry **2** with entries **22** and **25**).

The presence of a phenyl ring is important for the conversion to an ESC-like state, since the replacement with different heterocycles led to disparate reprogramming activity (entries **26-30**). On one hand, the introduction of a 1-naphthalenyl moiety as shown for derivative **26** (entry **26**) yielded a more active compound than **TR** (compare entry **2** with entry **26**). On the other hand, derivatives **27-30** (entries **27-30**) failed to deliver more potent agents compared to the parent hit, suggesting limited acceptance of heterocyclic moieties other than phenyl (entry **2**) and naphthyl (entry **26**).

In conclusion, the structure-activity relationship study within this series indicated that the range of allowed derivatizations at position C6 of the pteridine core is rather limited to just few key substitutions on the phenyl ring.

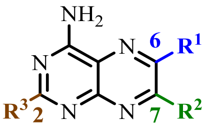
The next position addressed for derivatization was the amino group adjacent to the phenyl ring located at C7. Several derivatives were generated from commercially available derivative **31** following synthetic procedures already reported in the literature (see **Scheme 2**).[180] The biological activity of the newly generated compounds is shown in **Table 3**.



31 derivatives **32-34**

a. NaHCO$_3$ (2 eq), H$_2$O; then benzil (1 eq), reflux (3h) [derivatives **32** and **34**]
b. Phenylglyoxal (1.1 eq), hydroxylamine hydrochloride (1.1 eq), MeOH, reflux [derivative **33**]

Scheme 2. Synthesis route to access pteridine-derived compounds bearing substitutions at position C6 and C7.

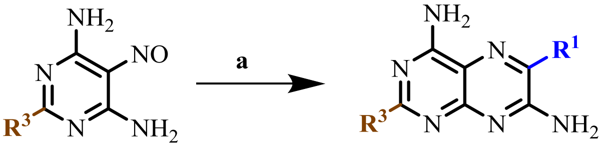


Nr.	R^1	R^2	R^3	Efficiency
2	-Ph	$-NH_2$	$-NH_2$	1
32	-Ph	-Ph	$-NH_2$	0.5
33	-Ph	-H	$-NH_2$	0.1
34	m-OH-Ph	m-OH-Ph	$-NH_2$	0.07

Table 3. Reprogramming efficiency of late-stage mEpiSCs upon treatment with derivatives **32-34** and **PD** at 10 µM concentration for 7 days. The relative reprogramming activity of **TR** (entry **2**) was arbitrary set to 1. **PD** = PD0325901 (MEK1 inhibitor). Data are representative of three independent experiments.

As shown in **Table 3**, the various substitutions at position C7 did not generate more active derivatives than the initial hit **TR** (compare entry **2** with entries **32-34**, **Table 3**). Whereas the introduction of a phenyl ring as shown for derivative **32** (entry **32**) diminished the activity to half, replacement of the original amino group with hydrogen (entry **33**) and 3'-hydroxyphenyl (entry **34**), respectively almost completely abolished the biological activity. These results suggest that the presence of the amino group at C7 is mandatory for an efficient reprogramming process.

The last derivatization investigated was at position C2 of the pteridine scaffold. Derivatives **37-40** were generated from the readily accessible starting materials **35** and **36**, respectively (see **Scheme 3**). Biological activities exhibited by these derivatives are presented in **Table 4**.

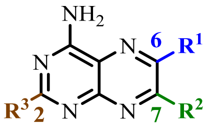


35 $R^3 = -CH_3$ $\quad$ $R^3 = -CH_3$ in derivatives **37-38**

36 $R^3 = -SCH_3$ $\quad$ $R^3 = -SCH_3$ in derivatives **39-40**

a. Phenylacetonitrile (1.1 eq), NaH (1.1 eq), 2-ethoxyethanol, reflux (1-3h)

Scheme 3. Synthesis route to access pteridine-derived compounds bearing substitutions at position C2 and C6.



Nr.	**R^1**	**R^2**	**R^3**	**Efficiency**
2	-Ph	-NH_2	-NH_2	1
37	-Ph	-NH_2	-CH_3	0
38	m-F-Ph	-NH_2	-CH_3	0
39	-Ph	-NH_2	-SCH_3	0.08
40	m-F-Ph	-NH_2	-SCH_3	1.72

Table 4. Reprogramming efficiency of late-stage mEpiSCs upon treatment with derivatives **37-40** and **PD** at 10 µM concentration for 7 days. The relative reprogramming activity of **TR** (entry **2**) was arbitrary set to 1. **PD** = PD0325901 (MEK1 inhibitor). Data are representative of three independent experiments.

As evidently shown in **Table 4**, the replacement of the initial amino group at C2 with either methyl or thiomethyl as shown for derivatives **37** and **39** (entries **37** and **39**) completely reduced the biological activity (compare entry **2** with entries **37** and **39**). This result indicates that the presence of the original amino group is imperative for the reprogramming process to occur. Taking in consideration the beneficial aspects upon introducing halogen atoms in the meta-position of the C6 phenyl ring, derivatives that maintain this substitution pattern and additionally possess methyl or thiomethyl groups in position C2 (entries **38** and **42**) were consequently synthesized. Only the addition of a fluorine atom in the case of thiomehyl derivative as shown by compound **40** (entry **40**) could reproduce the phenotype. Surprisingly, these compounds were even more active than **TR**, but less potent than the congeners containing the original amino group in position C2 (compare entry **13**, **Table 2** with entry **40**, **Table 4**). The current finding suggests that the amino-group at C2 position can be replaced with thiomethyl, however only in combination with fluorine in the meta-position of the phenyl ring located at C6.

In conclusion, the structure activity relationship indicated that the range of derivatizations permitted on the pteridine ring is rather limited (see **Figure 9**). Thus, only substitutions in the meta- and ortho-positions of the C6 phenyl ring could yield compounds with better reprogramming efficiencies than **TR**. Any tentative placement of groups other than hydrogen in the para position led to significant reduction or completely abolished the biological activity. Moreover, both amino groups in positions C2 and C7 were crucial, although the former could be replaced with thiomethyl moiety but only when the phenyl ring is equipped with a fluorine atom in the meta-position.

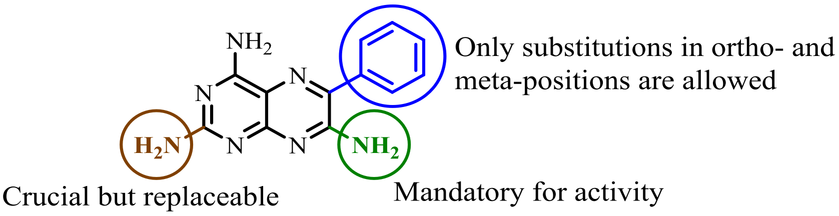
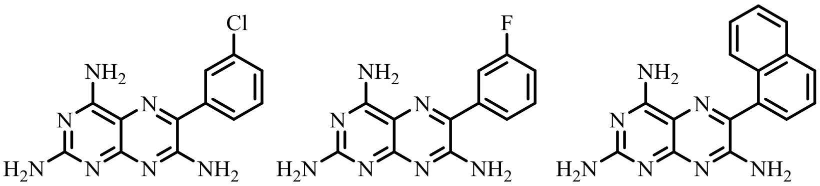


Figure 9. Conclusions of the structure-activity relationship study derived from the capacity to induce reprogramming of late-stage mEpiSCs by derivatives **2-40** in combination with **PD**, both compounds at 10 µM for 7 days. **PD** = PD0325901 (MEK1 inhibitor).

4.2. Characterization of the reprogramming activity

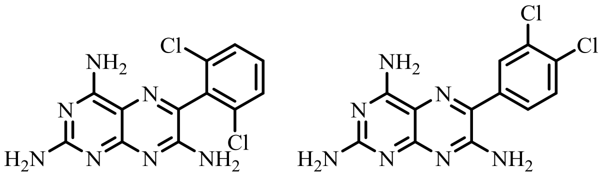
As shown in **Table 2-4**, several compounds were identified to be more potent compared to the parent derivative **TR**, namely derivatives **12**, **13**, **26**, **21** and **23** (entries **12**, **13**, **26**, **21** and **23**, **Table 2**).



Epiblastin A
(entry 12)
R.E. = 8.4

Epiblastin B
(entry 13)
R.E. = 3.1

Epiblastin C
(entry 26)
R.E. = 2.6



Epiblastin D
(entry 21)
R.E. = 2.2

Epiblastin E
(entry 23)
R.E. = 1.6

Figure 10. Chemical structures of Epiblastin A-E. The number between the brackets indicates the entry number shown in **Table 2**. R.E. is the abbreviation for Reprogramming Efficiency which in case of **TR** (entry **2**) was arbitrary set to 1.

As these compounds share the common ability of converting late-stage mEpiSCs into mES-like cells, these derivatives were designated Epiblastins, and were further labelled according to their decreasing biological activity as Epiblastin A-E, e.g. derivative **12** (entry **12**, **Table 2**) is termed Epiblastin A (see chemical structure in **Figure 10**). Due to their increased biological activity compared to **TR**, Epiblastin A-E were selected for further in-depth biological characterization. Epiblastin B contains a meta-fluoro substitution and is the second most active compound. The introduction of the 1-naphthalenyl moiety as in Epiblastin C slightly decreased the biological activity compared to Epiblastin B. Finally, the specific arrangement of two chlorine atoms on the phenyl ring in Epiblastin D and Epiblastin E, respectively, generated biologically more active compounds compared to **TR**.

Employing the same screening assay conditions established for **TR**[148,149] in 96-well format (experiments performed by Dr. Damir J. Illich, Max Plank Institute for Molecular Biomedicine, Münster, Germany), Epiblastin A-E derivatives were separately tested for reprogramming of late-stage mEpiSCs in lower throughput. **TR** was added in all the performed assays as the reference compound. Experiments carried out by Dr. Miao Zhang (Max Plank Institute for Molecular Biomedicine, Münster, Germany) confirmed that Epiblastin A-E reproducibly convert the late-stage mEpiSC to an Oct4-GFP-positive ESC-like population in combination with **PD**. As shown in **Figure 11**, these derivatives reproducibly achieve reprogramming efficiencies between 38 % and approximately 70 % as quantified by means of fluorescence activated cell sorting (FACS).

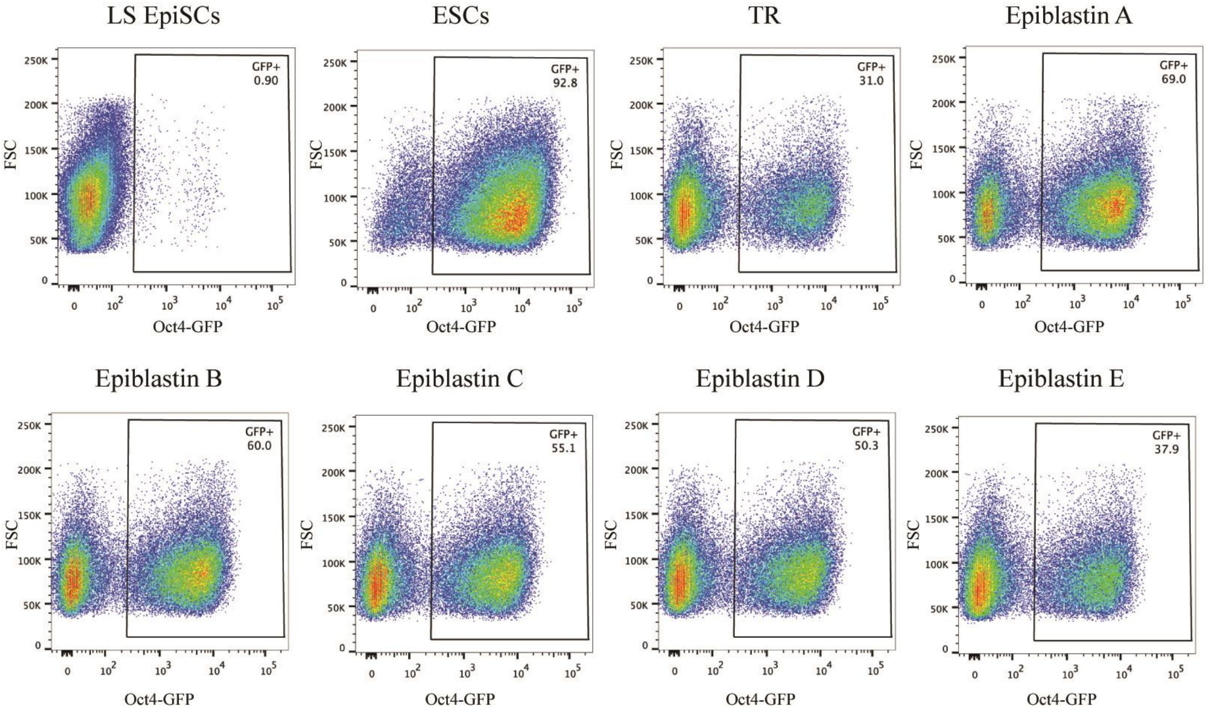


Figure 11. Reprogramming activity of **TR** and Epiblastin A-E on late-stage mEpiSCs evaluated by mean of the Oct4-GFP intensity of reprogrammed cells quantified using fluorescence-activated cell sorting (FACS). Reprogramming activity of **TR** / Epiblastin A-E has been measured in the presence PD0325901 (MEK1 inhibitor). All compounds were used at 10 μM concentration. LS EpiSCs = late-stage mouse epiblast stem cells; ESCs = mouse embryonic stem cells; **TR** = Triamterene. Representative histograms show the expression of Oct4-GFP (pluripotency-associated marker) in percentage on the x-axis. FSC = Forward scatter counts. Data are representative of three independent experiments.

Therefore, 38 % to 70 % of the initial cell population could reactivate the expression of *Oct4* upon treatment with Epiblastins A-E and **PD**. For comparison, the expression levels of the same gene construct in an authentic mESC line was evaluated to be approximately 93 %, whereas **TR** converted only 31 % of the starting population.

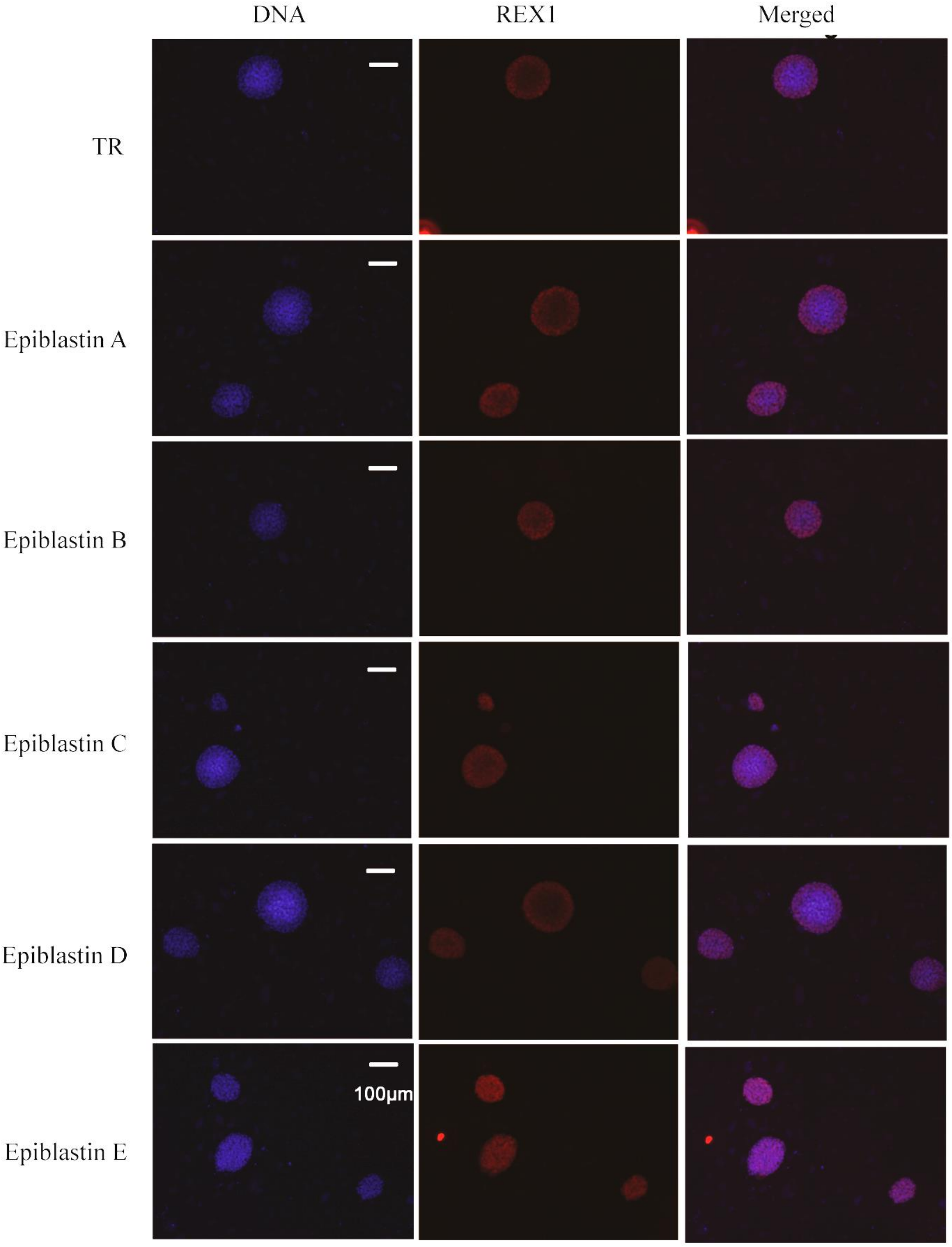
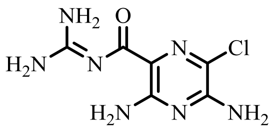

DNA
REX1
Merged
TR
Epiblastin A
Epiblastin B
Epiblastin C
Epiblastin D
Epiblastin E
100µm

Figure 12. Detection of REX1, an ESC-specific marker, in the reprogrammed cells generated upon treatment with **TR** / Epiblastin A-E in combination with **PD** at 10 µM for 7 days. **TR** = Triamterene; **PD** = PD0325901 (MEK1 inhibitor). Cells were stained for DNA (blue) and REX1 (red) using anti-Rex1 antibody and a secondary antibody coupled to Alexa Fluor® 568 and DAPI, respectively. Scale bar: 100 µm.

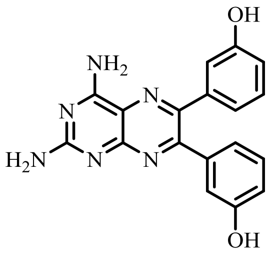
Immunofluorescence experiments showed that the Oct4-GFP-positive cell population obtained upon treatment with Epiblastin A-E and **PD** homogeneously express *Rex1* (also known as *Zpf42*),[128] a known marker associated with embryonic stem cells only (see **Figure 12**).[128,129] This result substantiates that the transition of the initial population indeed progressed to a naïve pluripotent state.

4.3. Pteridine-based analogs inhibit the CSNK1 kinase family

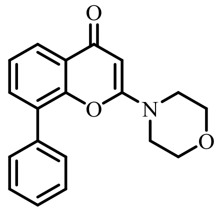
Since **TR** has been associated with antagonistic effects on the epithelial sodium channel (ENaC),[150] this target was first investigated to check if it is responsible for reprogramming late-stage mEpiSCs. The structurally unrelated ENaC inhibitor amiloride (see chemical structure in **Figure 13**) failed to reproduce the phenotype, indicating that ENaC is not responsible for the cellular transition. Additionally, a **TR**-related derivative termed TG100-115 (see chemical structure in **Figure 13**) has previously been reported to target various phosphatidylinositol 3-kinase (PI3K) isoforms with half-maximal inhibitory concentration values ranging from nanomolar to micromolar range.[196,197] Therefore, **TR** was been submitted to a kinase profiling study comprising more than 100 protein kinases that are a part of major signaling pathways governing various aspects of stem cell biology such as differentiation, self-renewal and reprogramming. The results of the screening campaign showed that **TR** is a fairly selective inhibitor of casein kinases 1 (CSNK1)[198,199] alpha 1 (thereafter A1), delta (thereafter D) and epsilon (thereafter E), in addition to PI3K isoforms.[148,149] The contribution of the PI3K family to the reprogramming mechanism was devalidated using structurally diverse inhibitors such as TG100-115, LY294002 and AS605240 (see chemical structures in **Figure 13**). Only upon treatment with the pan-CSNK1A1/D/E inhibitor D4476[200] (see chemical structure in **Figure 13**), the initial cell population reactivated Oct4-GFP expression.



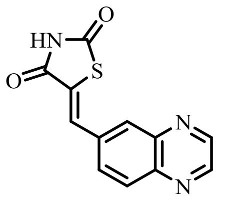
Amiloride
ENaC inhibitor



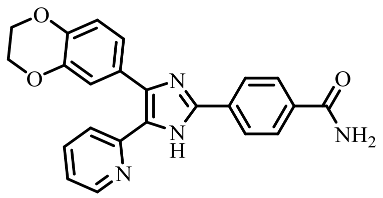
TG100-115
PI3K inhibitor



LY294002
PI3K inhibitor



AS605240
PI3K inhibitor



D4476
CSNK1 inhibitor

Figure 13. Chemical structures of the ENaC inhibitor amiloride, PI3K inhibitors TG100-115, LY294002 and AS605240 together with the pan-CSNK1A1/D/E inhibitor D4476.

By means of fluorescence-activated cell sorting (FACS), the biological action of the derivatives shown above was quantified. D4476 was found to be slightly more active than the initial found hit **TR**, suggesting that CSNK1 inhibition might, at least partially contribute to the reprogramming mechanism (see **Figure 14**).

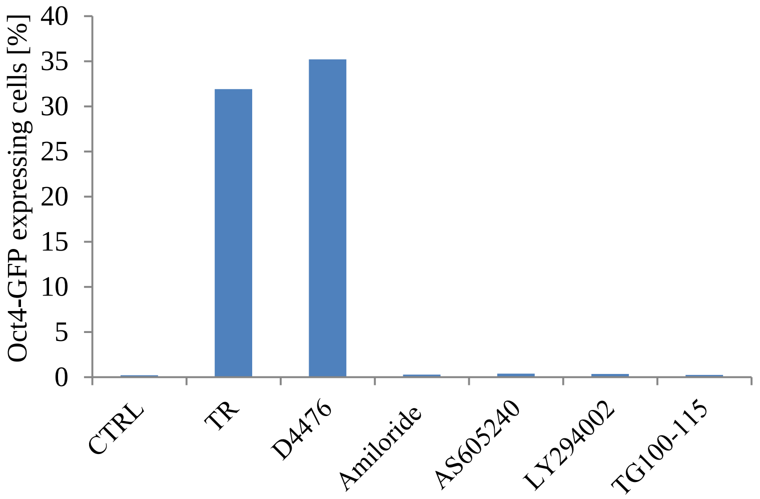


Figure 14. Reprogramming activity of ENaC inhibitor amiloride, PI3K inhibitors LY294002, AS605240 and TG100-115 together with the pan-CSNK1A1/D/E inhibitor D4476. CTRL = late-stage mEpiCSs treated with DMSO. Oct4-GFP intensity of reprogrammed cells was quantified by mean of fluorescence-activated cell sorting (FACS). Data are representative of three independent experiments.

Therefore, the pteridine-based compound collection was assayed against CSNK1D and CSNK1E isoenzymes by means of a luminescent-based readout using the KinaseGlo® reagent (Promega)[201] (see **Figure 15** for the chemical reaction underlying the signal generation).

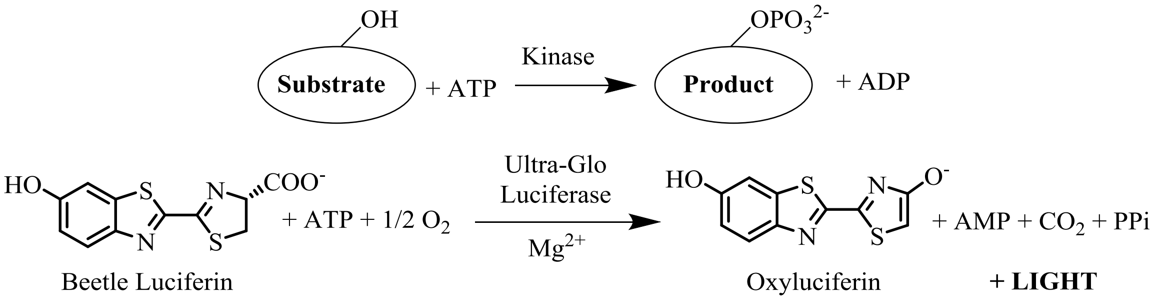


Figure 15. Chemical reactions within the KinaseGlo® assay. The activity of ATP-consuming enzymes (reactions) is detected by means of a decrease in the ATP concentration. Upon the ATPase (e.g. kinase) reaction, the remaining ATP is employed as substrate by the luciferase enzyme to catalyze the mono-oxygenation of luciferin to oxyluciferin producing one photon of light per turnover.

The principle of the KinaseGlo® reagent assay relies on the ability to evaluate the kinase activity by measuring the remaining ATP after the completion of the enzymatic reaction. Thus, the residual ATP is consumed in the presence of a thermostable luciferase (Ultra-Glo™) which uses Beetle Luciferin as substrate and molecular oxygen. As a consequence of this enzymatic reaction, a stable “glow-type” luminescent signal is generated, which is inversely related to kinase activity. One of the main advantages of the KinaseGlo® reagent is the possibility of measuring every type of kinase and substrate (peptide, protein, lipid, sugar) combination and its independency of any radiolabeled reagents. According to the manufacturer, the Ultra-Glo™ Luciferase is much less prone to small molecule interference than other ATP detection kits, thus reducing the false positive rate.[202,203] Many reports in the literature successfully used this platform for high throughput screening campaigns for other kinases[204,205,206] and even for CSNK1.[207,208,209]

His_6-tagged constructs of the human CSNK1D (amino acid sequence 1-294, thereafter designated CSNK1D) and CSNK1E (amino acid sequence 1-294, thereafter termed CSNK1E) lacking the C-terminus were designed in both cases. The sole purpose of generating constructs lacking the C-terminae domains was attributed to the fact that residues within this region have been previously reported to be phosphorylated by CSNK1 leading to subsequent autoinhibition of the kinase activity.[210,211,212] The assay conditions for the CSNK1D construct were optimized with respect to various parameters such as the concentration of the enzyme, substrate and ATP using various incubations and reaction times (initial assay optimization was carried out by Dr. Yasushi Takemoto, RIKEN-Max Planck Joint Research Center for Systems Chemical Biology, Wako, Japan; see **Figure 16**).

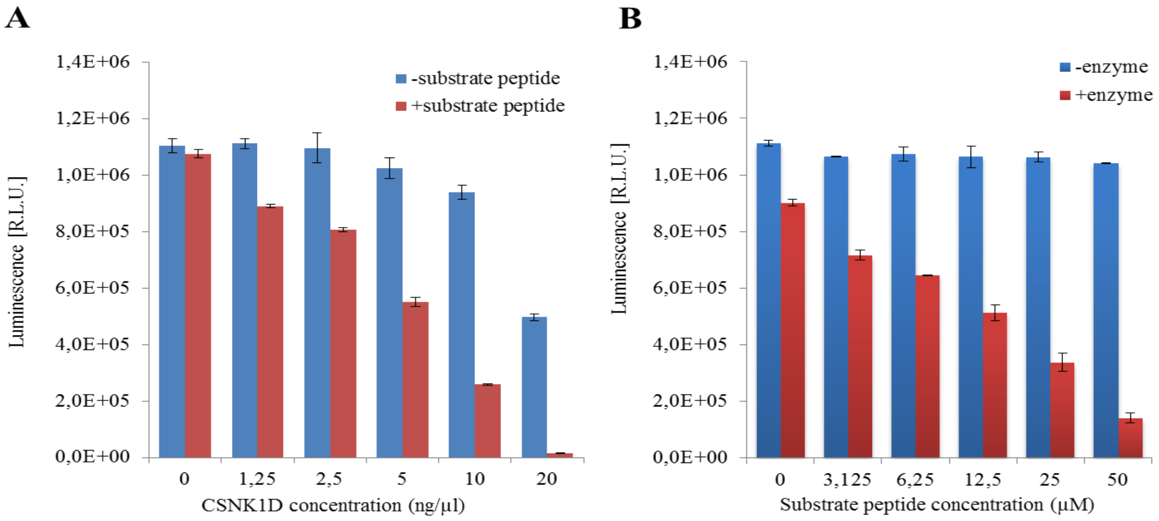


Figure 16. Optimization of the assay conditions with respect to the CSNK1D and substrate peptide concentration, respectively. (A) Determination of the optimal CSNK1D enzyme concentration at 50 µM substrate peptide concentration. Different concentrations of purified CSNK1D were incubated with and without substrate peptide. (B) Determination of the optimal substrate peptide concentration at a fixed CSNK1D enzyme concentration of 10 ng / µL. Different concentrations of substrate peptide were incubated with and without purified CSNK1D. The kinase activity was assessed by means of the KinaseGlo® reagent. The results represent mean values ± SD of three independent measurements performed in duplicate. [R.L.U. = Relative Light Units]

First, the optimal enzyme concentration was determined using 50 µM of the substrate peptide. As shown in **Figure 16A**, the best results were achieved by using 10 ng / µL of the CSNK1D kinase under the employed experimental conditions. In this case, the Z' factor[213] had an excellent value of 0.87.

After determining the suitable CSNK1D enzyme concentration for the assays, similar experiments were performed whereby the optimal concentration for the substrate peptide was evaluated. Accordingly, maintaining the CSNK1D enzyme concentration at 10 ng / µL, a broad range of substrate peptide concentrations was tested from 0 to 50 µM (see **Figure 16B**). The best results were obtained by using 50 µM of the substrate judging by the Z' factor which under these conditions was calculated to be 0.93. Gratifyingly, the conditions optimized for CSNK1D were successfully translated to the almost identical CSNK1E isoenzyme with very good Z' factor values of 0.9 and 0.7 for 10 ng / µL enzyme (see **Figure 17A**) and 50 µM substrate peptide (see **Figure 17B**) concentrations, respectively.

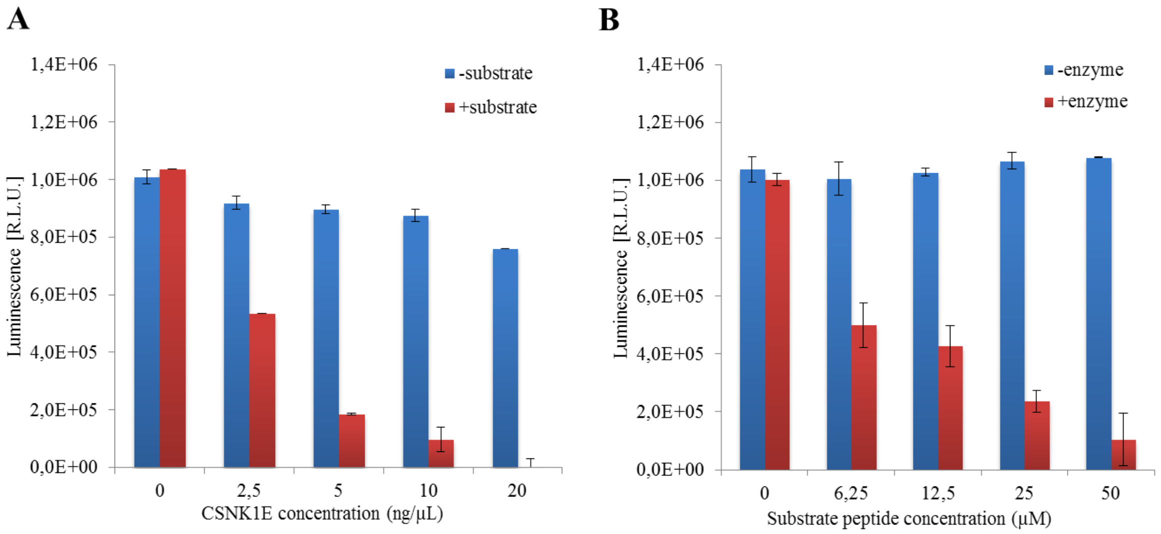


Figure 17. Optimization of the assay conditions with respect to the CSNK1E and substrate peptide concentration, respectively. (A) Determination of the optimal CSNK1E enzyme concentration at 50 µM substrate peptide concentration. Different concentrations of purified CSNK1E were incubated with and without substrate peptide. (B) Determination of the optimal substrate peptide concentration at a fixed CSNK1E enzyme concentration of 10 ng / µL. Different concentrations of substrate peptide were incubated with and without purified CSNK1E. The kinase activity was assessed by means of the KinaseGlo® reagent. The results represent mean values ± SD of three independent measurements performed in duplicate. [R.L.U. = Relative Light Units]

Furthermore the effect of DMSO on the CSNK1D enzymatic activity was determined. As a result, the final DMSO concentration was set at 1 % to counteract the high tendency for precipitation of the pteridine library members (data not shown).

With the optimized conditions in hands, the inhibitory activity for the previously published pteridine derivatives, **TR** and Epiblastin A, was evaluated for CSNK1D/E isoenzymes. The half-maximal inhibitory concentrations (IC_{50}) were found to be 36 µM and 0.5 µM, respectively against CSNK1D (see **Figure 18A**). These values are in good agreement with the previously published reports.[148,149] A good correlation was found in the case for CSNK1E inhibition by Epiblastin A as well, exhibiting an IC_{50} value of approximately 5 µM compared to 27 µM for **TR** (see **Figure 18B**). Furthermore, testing the structurally unrelated pan-CSNK1 inhibitor D4476[200] in the assay also gave a good correlation to the IC_{50} values previously reported by commercial profiling at Reaction Biology. Moreover, taking in consideration the broad range of IC_{50} values determined, the described assay can identify CSNK1D/E inhibitors with potencies (e.g., IC_{50} values) ranging from nanomolar to high micromolar range.

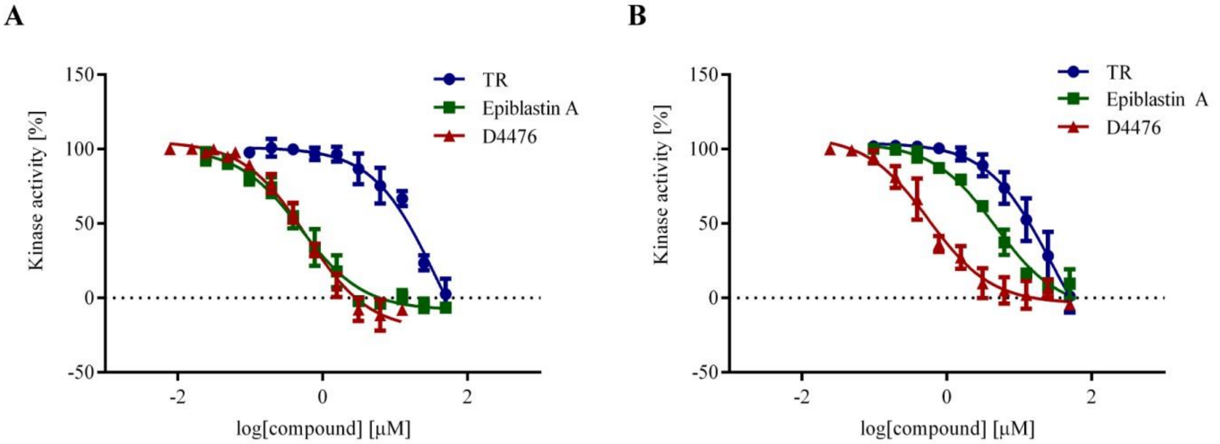


Figure 18. Dose-response curves of **TR**, Epiblastin A and D4476 for inhibition of CSNK1D (A) and CSNK1E (B). The curves represent the mean values ± SD of three independent measurements performed in duplicate. Determined IC_{50} value ± SD were generated by fitting the data with three parameter nonlinear regression analysis using GraphPad Prism version 6.00 for Windows.

Before proceeding with profiling the members of the library against CSNK1D/E isoenzymes, the mode of inhibition exhibited by the pteridine class of compounds was explored. This can be proven using the Lineweaver-Burk plot and the shift of the IC_{50} value upon modifying the ATP concentration in the assay. The Lineweaver-Burk plot is obtained by testing various concentrations of the inhibitor under investigation at different ATP concentrations. As a result, Epiblastin A was found to act as a competitive inhibitor of both CSNK1D (**Figure 19A**) and CSNK1E (**Figure 19B**) isoforms with respect to ATP.

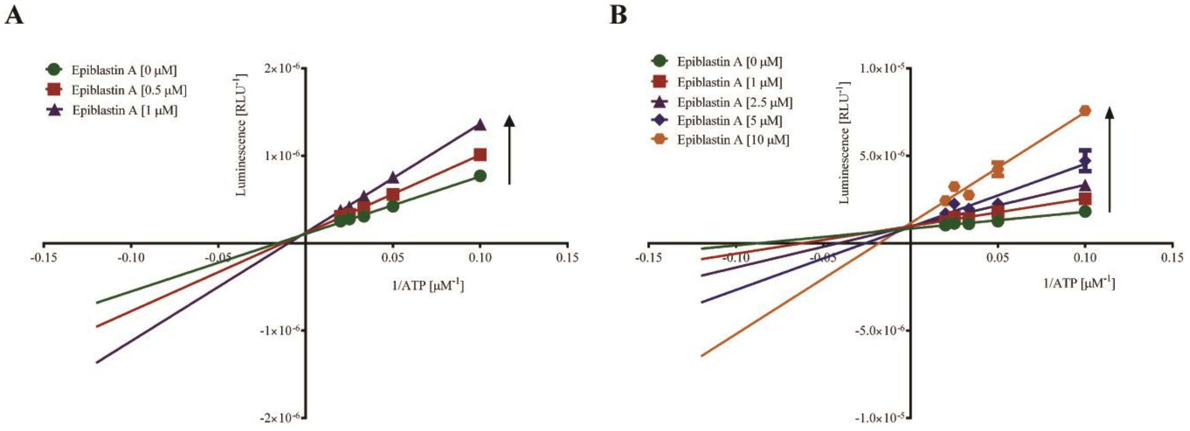


Figure 19. Determination of the CSNK1D/E kinase inhibition mode for Epiblastin A. Lineaweaver-Burk plot shows competitive inhibition of CSNK1D (A) and CSNK1E (B) by Epiblastin A with respect to various ATP concentrations. Data represent mean values ± SD of three independent measurements performed in duplicate. Black arrows indicate the increasing concentrations of Epiblastin A. [R.L.U. = Relative Light Units]

The ATP competitive nature of Epiblastin A inhibition of these isoenzymes was further proven by analyzing the IC_{50} value at 50 μM and 100 μM ATP (see **Figure 20A** for

CSNK1D and **Figure 20B** for CSNK1E). In this case, the IC_{50} value for CSNK1D/E shifts to higher value upon increasing ATP concentrations in the assay suggesting that Epiblastin A directly competes with ATP for the nucleotide binding pocket of CSNK1D/E. This type of inhibition is not surprising when considering that the 2,4-diaminopyrimidine unit mimics the structure of the naturally occurring purine moiety present in the ATP structure.

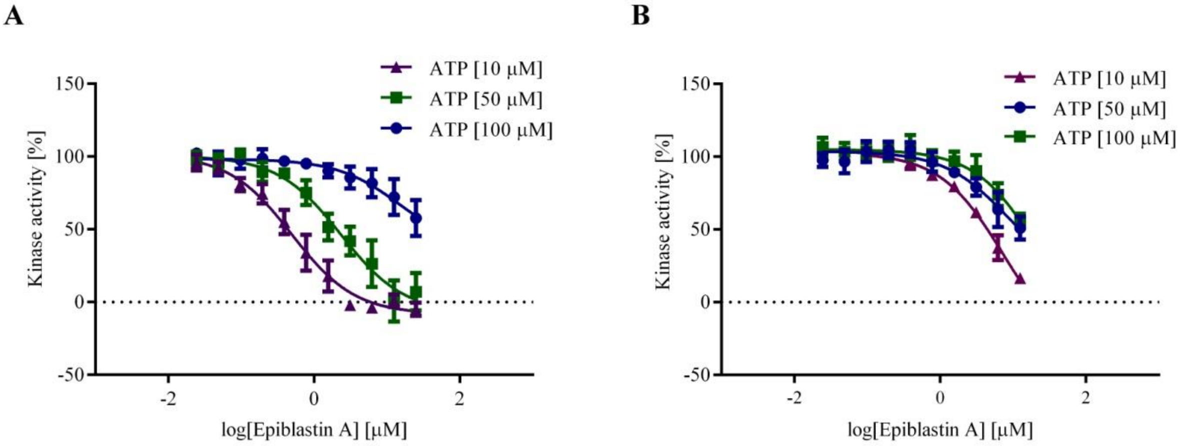


Figure 20. Half-maximal inhibitory determinations of Epiblastin A for inhibiting CSNK1D (A) and CSNK1E (B) at various ATP concentrations. Data represent mean values ± SD of three independent measurements performed in duplicate.

As a next step, the inhibition constant (K_i) of Epiblastin A for CSNK1D/E isoenzymes was determined by employing the Dixon plot at 10, 20 and 30 µM ATP concentration, respectively. The result of this analysis yielded for Epiblastin A an inhibition constant of 0.9 µM for CSNK1D (see **Figure 21A**) and 3.7 µM for CSNK1E (see **Figure 21B**).

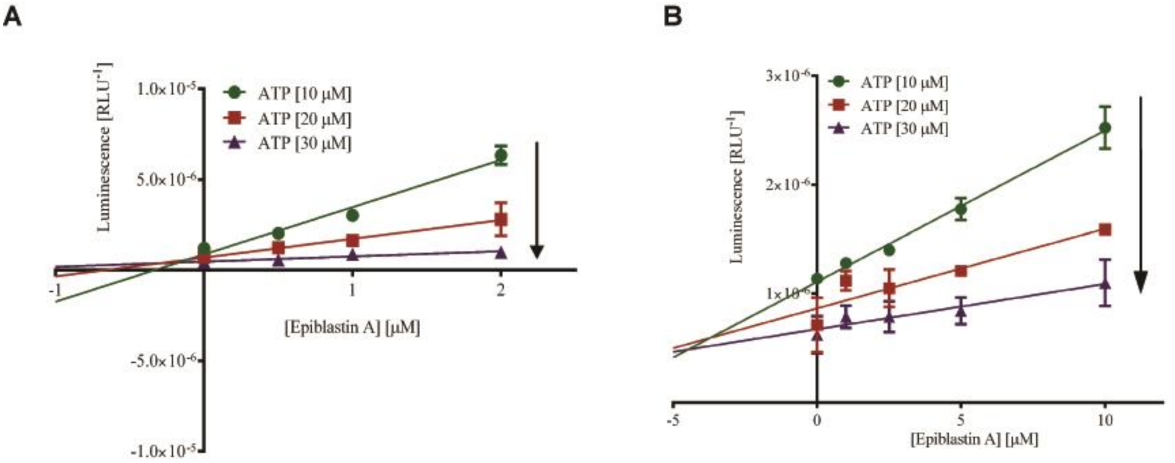


Figure 21. Dixon plot of Epiblastin A for inhibiting CSNK1D (A) and CSNK1E (B) yields inhibition constants (K_i) by using different concentrations of this derivative at various ATP concentrations. Data represent mean values ± SD of three independent measurements performed in duplicate. Black arrows indicate the increasing concentrations of ATP. [R.L.U. = Relative Light Units]

In summary, a luminescence-based assay using KinaseGlo® reagent characterized by very good Z' factor values was developed to monitor the enzymatic activity of CSNK1D and CSNK1E isoenzymes. Moreover, additional analyses disclosed a competitive type of inhibition of CSNK1D/E isoforms by Epiblastin A with respect to ATP. In the subsequent sections, the newly developed homogeneous assay was utilized to profile the pteridine library for identifying novel potent and / or selective CSNK1D/E inhibitors.

4.4. Screening of TR-based derivatives yields potent and / or selective CSNK1D/E inhibitors

Small changes on the phenyl ring attached to the pteridine core in position C6 are crucial for modulating the activity of various enzymes.[194,195,185] In consideration of these findings, the pteridine-based small molecule library was explored for potent and / or selective CSNK1D/E inhibitors. Before proceeding with assaying the enzymatic activity in the presence of the compounds, the interference of the pteridine derivatives with the luciferase reaction was crosschecked in absence of the kinase and the substrate to rule out false positives. Therefore, a single-point measurement study at the highest compound concentration employed in the assay, namely 50 µM, was employed. Fortunately, only a single compound (see entry 39, **Figure 22**) inhibited the luciferase enzymatic activity to various extents and was subsequently discarded from the study.

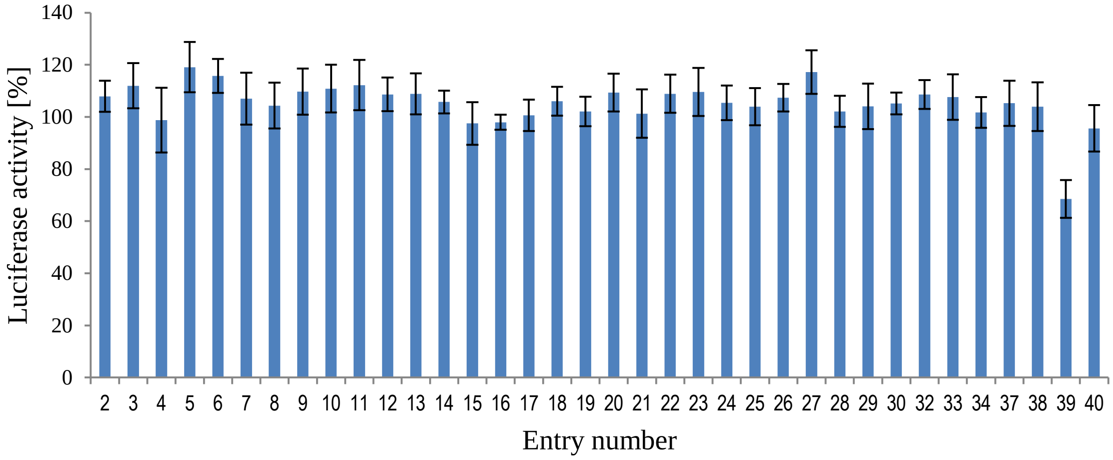


Figure 22. Single-point measurement analyses of pteridine derivatives **2-40** profiled at 50 µM for inhibition of the luciferin-luciferase background reaction. The results represent the mean values ± SD of three independent measurements performed in duplicate.

After having ruled out the interference by pteridine derivatives on luciferase reaction itself, the remaining derivatives were analyzed for inhibiting CSNK1D and CSNK1E isoenzymes at 50 μM compound concentration (see **Figure 23A** for CSNK1D and **Figure 23B** for CSNK1E).

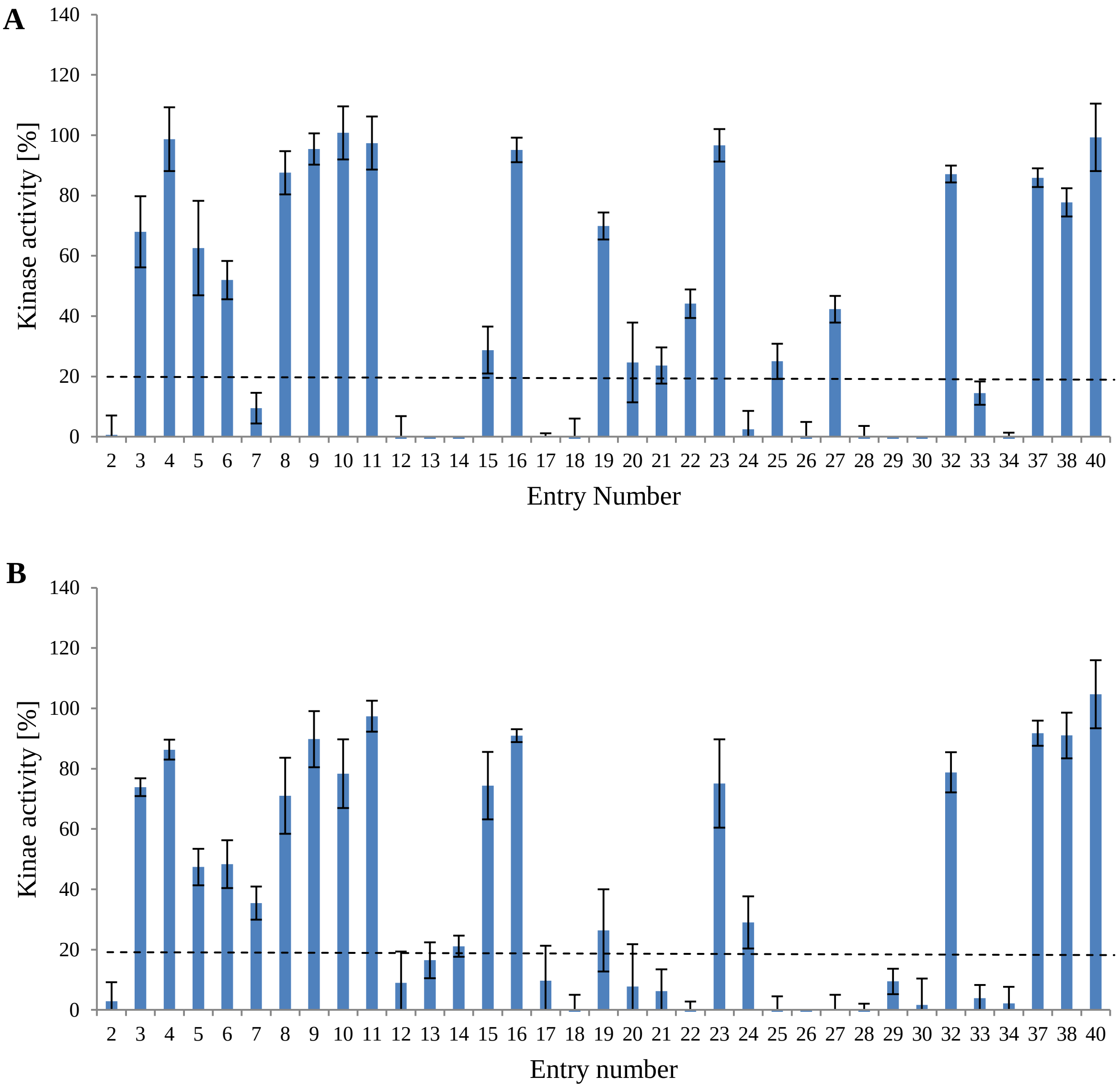
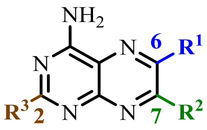


Figure 23. Single-point measurement analyses of pteridine derivatives **2-40** profiled at 50 μM concentration for inhibiting CSNK1D (A) and CSNK1E (B). The results represent the mean values ± SD of three independent measurements performed in duplicate. Black dashed line suggests the residual kinase activity threshold used for the compound selection. Derivatives below this line were further employed in concentration-dependent studies.

To generate an extended structure-activity relationship study, the compounds that showed at least 20 % residual kinase activity upon 50 μM compound treatment, i.e more than 80 % inhibition of the enzymatic reaction, were selected. All derivatives that fulfilled this criterion were further analyzed in dose-dependent experiments employing 10-point IC_{50} curves. The results of the screening campaign for CSNK1D/E are shown in **Table 5**. From these results, the following conclusions can be drawn (see entries **2-40**, **Table 5**).



Nr.	**R^1**	**R^2**	**R^3**	**IC_{50} [μM] CSNK1D**	**IC_{50} [μM] CSNK1E**
2	-Ph	$-NH_2$	$-NH_2$	36.4 ± 6.4	26.8 ± 4.2
3	p-F-Ph	$-NH_2$	$-NH_2$	68%	67%
4	p-OAllyl-Ph	$-NH_2$	$-NH_2$	99%	86%
5	p-OCH_3-Ph	$-NH_2$	$-NH_2$	63%	47%
6	p-CH_3-Ph	$-NH_2$	$-NH_2$	52%	48%
7	p-NH_2-Ph	$-NH_2$	$-NH_2$	15%	35%
8	p-NO_2-Ph	$-NH_2$	$-NH_2$	88%	71%
9	p-iPr-Ph	$-NH_2$	$-NH_2$	95%	90%
10	p-Ph-Ph	$-NH_2$	$-NH_2$	101%	78%
11	p-Obz-Ph	$-NH_2$	$-NH_2$	97%	97%
12	m-Cl-Ph	$-NH_2$	$-NH_2$	0.5 ± 0.1	4.7 ± 0.8
13	m-F-Ph	$-NH_2$	$-NH_2$	61.8 ± 12.5	14.4 ± 4.7
14	m-Me-Ph	$-NH_2$	$-NH_2$	20.2 ± 4.7	30%**
15	m-Br-Ph	$-NH_2$	$-NH_2$	2.5 ± 0.5	42%*
16	m-CF_3-Ph	$-NH_2$	$-NH_2$	95%	91%
17	m-OCH_3-Ph	$-NH_2$	$-NH_2$	12.3 ± 1.9	26.3 ± 4.9
18	m-NH_2-Ph	$-NH_2$	$-NH_2$	32.4 ± 7.0	12.4 ± 3.0
19	o-OCH_3-Ph	$-NH_2$	$-NH_2$	70%	26%
20	o-F-Ph	$-NH_2$	$-NH_2$	25%**	28.3 ± 3.8
21	o,o‘-diCl-Ph	$-NH_2$	$-NH_2$	24%**	38.9 ± 5.8
22	o,p-diCl-Ph	$-NH_2$	$-NH_2$	44%**	4.3 ± 0.9
23	m,p-diCl-Ph	$-NH_2$	$-NH_2$	97%	75%
24	o,o‘-diF-Ph	$-NH_2$	$-NH_2$	35.7 ± 5.2	51.5 ± 11.8
25	m,o‘-diF-Ph	$-NH_2$	$-NH_2$	25%**	38.4 ± 6.1
26	Naphtalen-1-yl	$-NH_2$	$-NH_2$	36.5 ± 6.4	1.0 ± 0.2
27	Pyridin-2-yl	$-NH_2$	$-NH_2$	42%*	6.6 ± 1.9
28	Benzo[b] thiophen-3-yl	$-NH_2$	$-NH_2$	2.8 ± 0.4	0.6 ± 0.1
29	Thiophen-2-yl	$-NH_2$	$-NH_2$	24.4 ± 6.4	8.1 ± 2.5
30	Pyridin-3-yl	$-NH_2$	$-NH_2$	16.1 ± 4.3	15.5 ± 4.0
32	-Ph	-Ph	$-NH_2$	87%	79%
33	-Ph	-H	$-NH_2$	80.6 ± 11.7	6.0 ± 1.5
34	m-OH-Ph	m-OH-Ph	$-NH_2$	9.6 ± 1.6	6.6 ± 1.5
37	-Ph	$-NH_2$	$-CH_3$	86%	92%
38	m-F-Ph	$-NH_2$	$-CH_3$	78%	91%
40	m-F-Ph	$-NH_2$	$-SCH_3$	99%	105%

Table 5. Activity of pteridine derivatives **2-40** for inhibition of CSNK1D and CSNK1E. Data represent mean values ± SD of three independent measurements performed in duplicate. Residual activities were determined upon incubation with 50 μM pteridine derivative; * = residual activity at 12.5 μM, higher concentration lead to precipitation; ** = the IC_{50} curve has not reach saturation and thus no proper IC_{50} value could be determined. Determined IC_{50} value ± SD were generated by fitting the data with three parameter nonlinear regression analysis using GraphPad Prism version 6.00 for Windows.

Compared to the unsubstituted derivative, **TR** (entry **2**, IC_{50} = 36.4 ± 6.4 μM for CSNK1D and IC_{50} = 26.3 ± 4.9 μM for CSNK1E, respectively), any other derivatization of the para-position of the phenyl ring is detrimental for the inhibition of both CSNK1D/E isoenzymes.

Even a fluorine atom as shown by derivative **3** (entry **3**) yielded an inactive compound exhibiting 68 % and 67 % residual kinase activity for CSNK1D and CSNK1E, respectively when treated with 50 μM derivative concentration. Any other substitutions lead to an almost complete loss of the enzymatic activity suggesting that this position is not amenable for substitution (entries **4-11**). However, derivatization of the meta-position proved to be more accessible (entries **12-18**). For example, the most active inhibitor against CSNK1D identified so far is Epiblastin A (entry **12**, IC_{50} = 0.5 ± 0.1 μM) which features a chlorine atom in the meta-position. Interestingly, the replacement by a bromine atom as shown for derivative **15** (entry **15**) did not dramatically reduced the activity against the same isoform (IC_{50} = 2.5 ± 0.5 μM), while the fluorine atom in Epiblastin B (entry **13**) yielded a relatively inactive derivative (IC_{50} = 61.8 ± 12.5 μM). Of side note, the iodo-derivative was also synthesized but could not be used in the study due to severe solubility problems. The halogenated series clearly shows that the meta-chloro derivatization fulfills best the steric and electronic requirements for achieving a high inhibition degree against CSNK1D (see section *4.6. Crystallization studies*). Additionally, the methoxy moiety was also tested, as well as methyl- and amino-substitutions in the same position in derivatives **17**, **14** and **18**, respectively (entries **17**, **14** and **18**). The compounds exhibited IC_{50} values of 12, 20 and 32 μM, respectively. These decorations are electron-donating groups in nature and might imply a possible influence of the electron density of the phenyl ring on the inhibition process.

The situation is different when analyzing the impact of the meta-substitution on the closely related CSNK1E isoform. The introduction of a chlorine atom as shown for Epiblastin A (entry **12**) generated an IC_{50} value of 4.7 μM, which is approximately ten times less active than in the case of CSNK1D. In the same time, meta-fluoro substituted compound **13** (entry **13**) exhibited an IC_{50} for CSNK1E of 14 μM, while the bromine substitution as in derivative **15** (entry **15**) could not completely inhibit CSNK1E activity and reduced the residual kinase activity to 58 % at a concentration of 12.5 μM. Higher concentrations for this derivative lead to precipitation. Worthwhile are the meta-amino and meta-methoxy derivatives **18** and **17** (entry **18** and entry **17**) that showed IC_{50} values for CSNK1E of 12 μM and 25 μM, respectively. Moreover, 50 μM of the methyl-substituted compound **14** (entry **14**) could not entirely inhibit the CSNK1E enzymatic activity, i.e. at this concentration the residual kinase activity was 30 %. For both the CSNK1D/E isoenzymes, the trifluoromethyl containing derivative **16** (entry **16**) was found to be completely inactive. A limited number of compounds bearing ortho-substitution on the C6 phenyl ring did not yield potent derivatives against neither CSNK1D nor CSNK1E.

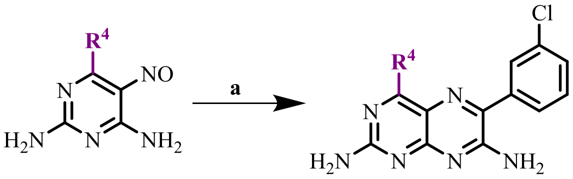
In conclusion, decorating the meta-position of the phenyl ring with either a chlorine- or bromine-atom as shown for Epiblastin A and derivative **15** (entries **12** and **15**) represents a promising strategy to obtain potent CSNK1D inhibitors. For example, Epiblastin A shows a good selectivity profile for CSNK1D, outmatching other commercial available CSNK1D inhibitors such as PF-670462.[214] Furthermore, the compounds reported here are excellent starting points for medicinal chemistry campaigns directed towards improving the activity and / or selectivity profile towards CSNK1D. The current findings highlight the importance of the substitution pattern on the phenyl ring at position C6 to achieve high potency and / or selectivity, which is also in line with previous reports for other classes of enzymes.

Disubstitution on the phenyl ring furnished interesting features of the pteridine derivatives (entries **21-25**, **Table 5**). For example, the chlorine atom distribution dramatically impacts the inhibitory activity especially in the case of CSNK1E. Thus, simultaneous introduction of a chlorine atom in ortho- and para-positions as shown in the case of derivative **22** (entry **22**) yielded the most active inhibitor of this series with an estimated IC_{50} value of 4.3 μM for CSNK1E. As a comparison, CSNK1D isoenzyme exhibited 44 % residual enzymatic activity upon treatment with this inhibitor at a concentration of 50 μM. Other chlorine substitutions on the phenyl ring yielded inactive derivatives for both CSNK1D/E. Similar results were obtained with a limited number of compounds bearing fluorine derivatizations in various positions on the phenyl ring at C6. In conclusion, decorating the aromatic moiety with two chlorine atoms in specific positions could represent the basis of generating selective CSNK1E inhibitors.

Startling results were obtained when the phenyl ring was replaced by various heterocycles (entries **26-30**, **Table 5**). For example, 2-thiophenyl substitution as shown in the case of derivative **29** (entry **29**) yielded an inhibitor with increased potency compared to **TR** (entry **2**) especially for CSNK1E ($IC_{50} = 24.4 \pm 6.4$ μM for CSNK1D and $IC_{50} = 8.1 \pm 2.5$ μM for CSNK1E). 3-Pyridinyl substitution as shown for derivative **30** (entry **30**) yielded a more active agent than **TR** (entry **2**) with respect to both CSNK1 isoforms ($IC_{50} = 16.1 \pm 4.3$ μM for CSNK1D and 15.5 ± 4.0 μM for CSNK1E, respectively). Moreover, 2-pyridinyl compound **27** (entry **27**) showed some degree of selectivity towards CSNK1E (IC_{50} = 6.6 ± 1.9 μM for CSNK1E compared to 42 % residual kinase activity for CSNK1D at a concentration of 50 μM). Thus, it might be possible that the nitrogen atom on the pyridine unit forms an intramolecular hydrogen-bond with the amino group in position C7 and enables trapping of the phenyl ring in a specific conformation that is selectively blocking the CSNK1E enzymatic activity. The replacement of the original phenyl ring at C6 with a 1-naphtalenyl moiety yielded Epiblastin C (entry **26**), one of the most selective CSNK1E inhibitors to date (IC_{50} = 36.5 ± 6.4 μM for CSNK1D and IC_{50} = 1.0 ± 0.2 μM for CSNK1E). Similar selectivity profile was reported for PF-4800567[214] with IC_{50} values of 32 and 711 nM for CSNK1E and CSNK1D, respectively. These findings imply that further optimization of Epiblastin C core structure could substantially improve both the potency and selectivity towards CSNK1E. Derivative **28** (entry **28**), containing the benzo[b]thiophen-3-yl group at the C6 yielded the most potent CSNK1E inhibitor across this series, with an estimated half-maximal inhibitory concentration value in the low micromolar range, i.e. IC_{50} = 0.6 ± 0.1 μM (IC_{50} = 2.8 ± 0.4 μM for CSNK1D). These values indicate a completely opposite inhibitory profile compared to Epiblastin A.

Finally, attempts were made to gain further insights from modifying other positions on the pteridine ring (entries **32-34**, **Table 5**). Removing the amino group in position C7 as shown for derivative **33** (entry **33**) yielded yet another selective CSNK1E inhibitor with an estimated IC_{50} value of 6 μM (IC_{50} = 80.6 ± 11.7 μM for CSNK1D). Any other substitution of the amino group at the position C7 yielded inactive derivative (see entry **32**). One interesting example is compound **34** (entry **34**) known in the literature as TG100-115,[197] bearing the 3,3'-dihydroxyphenyl substitutions in positions C6 and C7, respectively. As previously mentioned, this compound is reported as a potent class I PI3K inhibitor[180] and is currently in clinical trials for myocardial infarct.[215] Besides the PI3K family, TG100-115 also targets CSNK1D and CSNK1E isoenzymes with IC_{50} values of 10 μM and 7 μM, respectively. In conclusion, the amino group located at position C7 is critical for inhibiting CSNK1D/E isoforms. However, the replacement with a simple hydrogen atom confers a certain degree of selectivity towards CSNK1E.

At this point, the impact of derivatization at position C2 was investigated as well. Towards this end, the original amino group was replaced with either a methyl or thiomethyl group (entries **37-40**, **Table 5**). These substitutions yielded inactive compounds of both CSNK1D/E isoforms, highlighting the crucial role of the amino group in position C2 for the inhibition process. Even the addition of a fluorine or chlorine atom that proved beneficial for increasing the inhibitory activity towards CSNK1D could not block the enzymatic activity of CSNK1D/E isoenzymes. Motivated by these findings, a small number of derivatives was synthesized that contain an alkoxy substituent in position C4 on the pteridine ring but kept the meta-chloro derivatization at position C6. The rationale behind generating these compounds was to investigate if these derivatives are more active CSNK1D inhibitors compared to Epiblastin A. Synthesis of 4-alkoxy-2,7-diamino-6(3-chlorophenyl) pteridines was straightforward from readily accessible starting materials **41-49** as shown in **Scheme 4**.



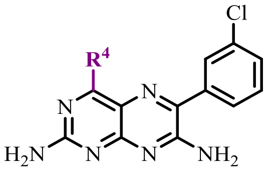
R^4 = alkyl, benzyl
in derivatives **41-49**

R^4 = alkyl, benzyl
in derivatives **50-59**

a. (3-chlorophenyl)acetonitrile (1.2 eq), Cs_2CO_3 (1.2 eq), THF, 65°C, 3h

Scheme 4. Synthesis of 4-alkoxy-2,7-diamino-6(3-chlorophenyl)pteridines **50-59**.

Inhibition studies performed with CSNK1D were performed by Dr. Yasushi Takemoto (RIKEN-Max Planck Joint Research Center for Systems Chemical Biology, Wako, Japan) and are shown in **Table 6**.



Nr.	**R^4**	**CSNK1D activity upon 1 µM Epiblastin A**	**CSNK1D activity upon 10 µM Epiblastin A**
2	$-NH_2$	17%	1%
50	ethyloxy	101%	91%
51	n-propyloxy	102%	100%
52	isopropyloxy	105%	85%
53	allyloxy	106%	93%
54	n-butyloxy	103%	93%
55	sec-butyloxy	102%	98%
56	cyclopentyloxy	102%	99%
57	cyclopentylmethyloxy	102%	101%
58	cyclohexylmethyloxy	101%	97%
59	benzyloxy	105%	101%

Table 6. CSNK1D residual enzymatic activity upon treatment with Epiblastin A-based derivatives **50-59** at 1 µM and 10 µM, respectively. Data represent the mean values of three independent measurements performed in duplicate.

Compared to Epiblastin A, the newly synthesized derivatives bearing alkoxy groups at the position C4 failed to inhibit CSNK1D enzymatic activity upon 1 μM and 10 μM compound treatments, respectively (compare entry **2** with entries **50-59**). These results suggest that the amino group in position C4 is imperative for exhibiting the inhibitory activity against CSNK1D at least.

In conclusion, the structure-activity relationships with respect to the inhibition of CSNK1D/E isoenzymes highlight the features also emphasized in **Figure 24**: (1) substitutions at C6 generate potent and / or selective CSNK1D/E inhibitors; (2) the amino group at position C7 is mandatory for maintaining the inhibitory activity against CSNK1D/E; replacement with a simple hydrogen atom renders selectivity towards CSNK1E; (3) amino groups at C2 and C4 are crucial for inhibiting CSNK1D/E; it can be surmised that these amino groups engage H-bond donors inside the ATP binding pocket of both kinases. Therefore, a defined array of H-bond donors and acceptors might be essential for exhibiting proper inhibitory activity.

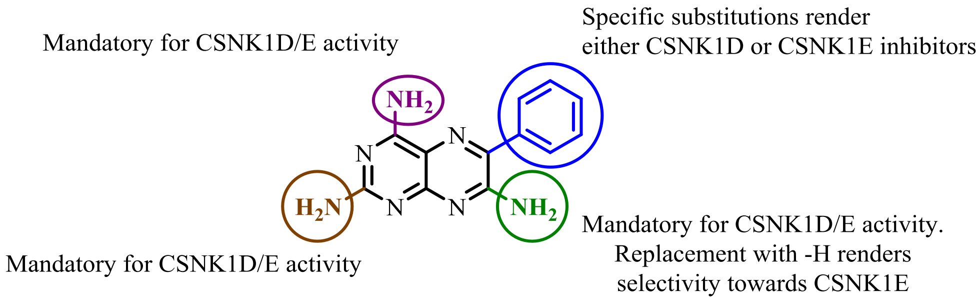
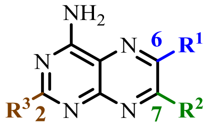

Mandatory for CSNK1D/E activity
Specific substitutions render either CSNK1D or CSNK1E inhibitors
NH_2
N
N
N
N
H_2N
NH_2
Mandatory for CSNK1D/E activity
Mandatory for CSNK1D/E activity. Replacement with -H renders selectivity towards CSNK1E

Figure 24. Conclusions of the structure-activity relationship study derived upon treatment of CSNK1D/E with pteridine derivatives **2-40** and **50-59**, respectively.

4.5. Evaluation of the selected pteridine analogues against CSNK1A1

Encouraged by the good inhibition profiles obtained for CSNK1D/E, several derivatives were tested for inhibition of CSNK1A1, a closely related CSNK1D/E isoform. Because of difficulties encountered with the expression of this particular CSNK1 isoenzyme, the selected compounds were profiled externally at Reaction Biology. Concentration-dependent measurements indicated that the selected compounds are fairly inactive for inhibition of CSNK1A1 isoform (see **Table 7**). As shown in **Table 7**, the manipulation of the substitution pattern of the pteridine ring in various positions allowed the generation of potent and / or selective CSNK1A1/D/E inhibitors. Thus, meta-substitution with a chlorine atom as shown for Epiblastin A (entry **12**) yielded a selective CSNK1D inhibitor with respect to both CSNK1E and CSNK1A1.



Entry	**R^1**	**R^2**	**R^3**	**IC$_{50}$ CSNK1A1 [µM]**	**IC$_{50}$ CSNK1D [µM]**	**IC$_{50}$ CSNK1E [µM]**
2	-Ph	-NH$_2$	-NH$_2$	36.81 ± 5.57	36.44 ± 6.49	26.82 ± 4.20
12	m-Cl-Ph	-NH$_2$	-NH$_2$	8.91 ± 2.38	0.50 ± 0.09	4.72 ± 0.79
15	m-Br-Ph	-NH$_2$	-NH$_2$	9.32 ± 1.88	2.46 ± 0.54	42%***
13	m-F-Ph	-NH$_2$	-NH$_2$	33.29 ± 4.89	61.79 ± 12.54	14.42 ± 4.70
22	o,p-diCl-Ph	-NH$_2$	-NH$_2$	55.56 ± 8.22	44%**	4.30 ± 0.88
26	Naphtalen-1-yl	-NH$_2$	-NH$_2$	35.94 ± 6.63	36.49 ± 6.42	1.00 ± 0.21
27	Pyridin-2-yl	-NH$_2$	-NH$_2$	54%*	42%**	6.59 ± 1.95
28	Benzo[b]thiophen-3-yl	-NH$_2$	-NH$_2$	5.61 ± 1.13	2.77 ± 0.39	0.56 ± 0.11
29	Thiophen-2-yl	-NH$_2$	-NH$_2$	14.23 ± 2.96	24.44 ± 6.37	8.15 ± 2.47
33	Ph	H	-NH$_2$	29.46 ± 5.02	80.61 ± 11.66	5.99 ± 1.54
34	m-OH-Ph	m-OH-Ph	-NH$_2$	5.52 ± 1.30	9.65 ± 1.63	6.59 ± 1.51

Table 7. IC$_{50}$ values for inhibition of CSNK1A1/D/E by a selected set of pteridine derivatives. Data represent mean values ± SD of three independent measurements for CSNK1A and performed in duplicate for CSNK1D/E. Determined IC$_{50}$ value ± SD were generated by fitting the data with three parameter nonlinear regression analysis using GraphPad Prism version 6.00 for Windows. Where no IC$_{50}$ determination was possible, the residual activities are shown as follows: * = residual activity upon treatment with 100 µM compound; ** = residual activity upon treatment with 50 µM compound; *** = residual activity upon treatment with 12.5 µM compound.

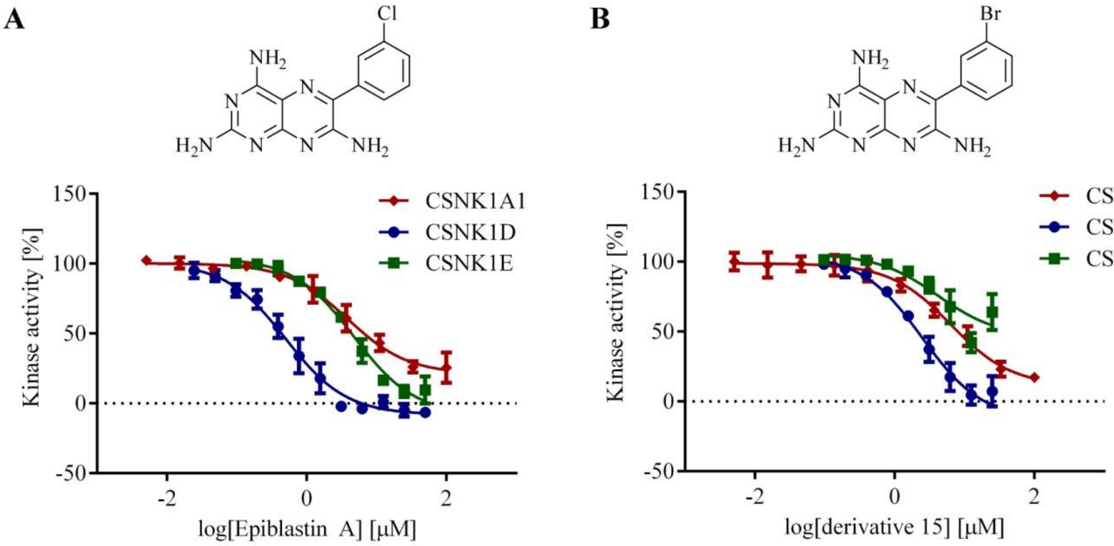


Figure 25. Dose-dependent profiles of Epiblastin A (A) and derivative **15** (B) for inhibition of CSNK1A1/D/E. Chemical structures for these compounds are highlighted in the upper part of each graph. Data represent mean values ± SD of three independent measurements performed in duplicate. Determined IC$_{50}$ value ± SD were generated by fitting the data with three parameter nonlinear regression analysis using GraphPad Prism version 6.00 for Windows.

Interestingly, a bromine atom as shown for derivative **15** (entry **15**) in the same position furnished a dual CSNK1A1/D inhibitor. The inhibition profiles of Epiblastin A and derivative **15** (entry **15**) for CSNK1A1/D/E isoenzymes are shown in **Figure 25,** together with the corresponding chemical structure.

Conversely, Epiblastin C and derivatives **22**, **27** and **33** (entries **26**, **22**, **27** and **33**) rendered selective CSNK1E inhibitors (see **Figure 26**).

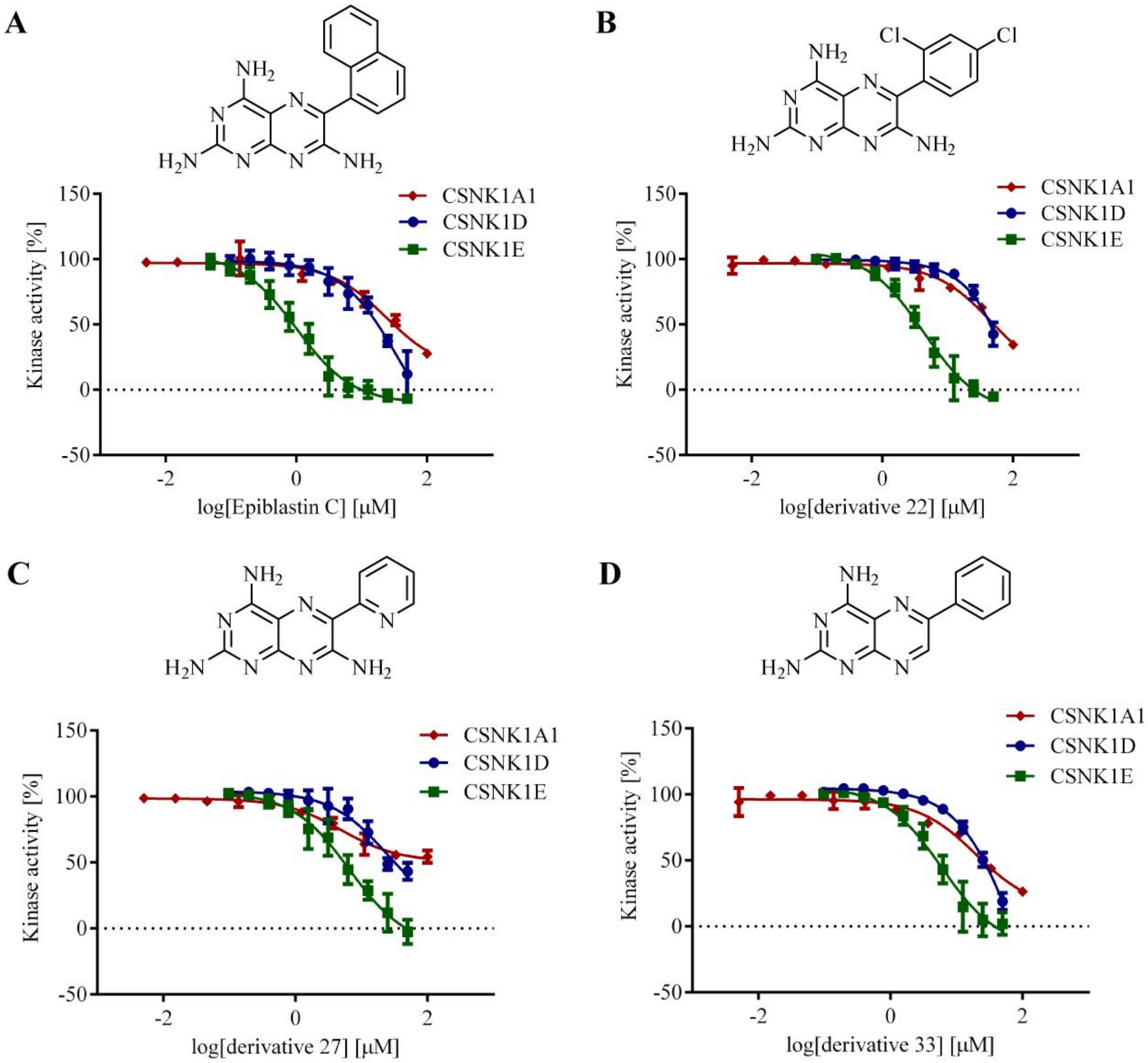


Figure 26. Dose-dependent profiles of Epiblastin C (A) and derivatives **22** (B), **27** (C), and **33** (D) for inhibition of CSNK1A1/D/E. Chemical structures for these compounds are highlighted in the upper part of each graph. Data represent mean values ± SD of three independent measurements performed in duplicate. Determined IC_{50} value ± SD were generated by fitting the data with three parameter nonlinear regression analysis using GraphPad Prism version 6.00 for Windows.

As shown above, Epiblastin A and Epiblastin C are fairly selective CSNK1D/CSNK1E inhibitors, respectively that open up new avenues for studying CSNK1-modulated processes in an isoform-specific fashion. Moreover, they represent good starting candidates for further optimization.

4.6. Crystallization studies

Crystallization studies were performed in collaboration with Arthur Porfetye (Max Planck Institute of Molecular Physiology, Dortmund, Germany). A crystal structure of the apo-CSNK1D (amino acid sequence 1-294) was obtained and resolved at a 1.9 Å resolution (with an R_{free} of 25.3 %) using synchrotron radiation at PXII. The structure of the apo-CSNK1D isoenzyme is depicted in **Figure 27**.

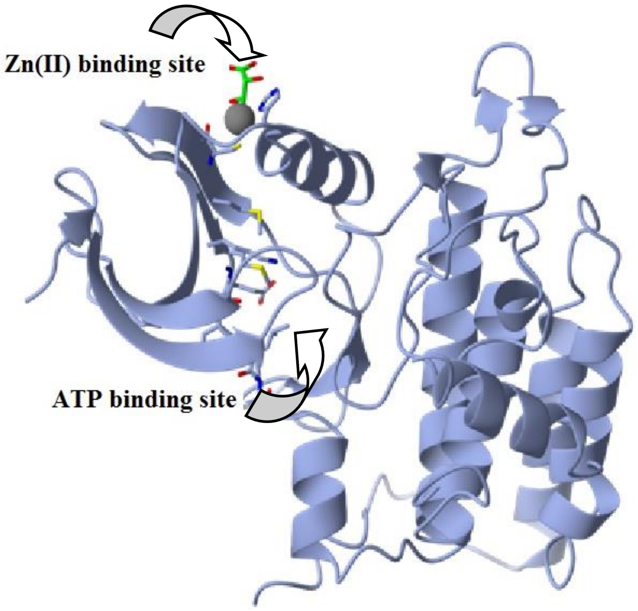


Figure 27. Cartoon representation of the apo-CSNKD1 kinase crystal structure, emphasizing the amino acid residues involved in ATP and Zn^{2+} binding, respectively.

The CSNK1D (amino acids 1-294) kinase construct has been previously employed by other groups for crystallization campaigns.[216] Under the used experimental conditions, the enzyme crystallized in the $P3_121$ crystal form, containing 1 protein molecule per asymmetric unit. This is different from the previously reported CSNK1D crystal structure (PDB ID: 3UYS),[217] which was solved in the P1 form containing 4 kinase monomers per asymmetric unit.

A notable exception in the crystal structure determined in the current study is the presence of a Zn^{2+} ion coordinated by the amino acid side chains His50, Cys71 and His278 of the symmetry mate and (*S*,*R*)-tartrate from the crystallization buffer (see **Figure 27**). Additionally, the nucleotide binding pocket of the CSNK1D unit is unoccupied, in contrast to the previously characterized structure, where this region was sterically blocked by the Arg16 amino acid side chain of another CSNK1D unit present in the crystal form.

An overlay of the generated apo-CSNK1D structure with monomer C of the prior report (PDB ID: 3UYS) indicates no significant differences, as shown in **Figure 28** (0.78 Å over 281 amino acids).

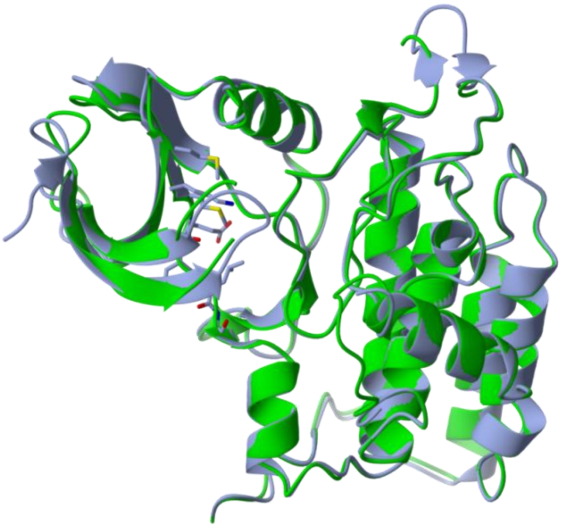


Figure 28. Structural overlap of the CSNK1D crystal structures obtained in the current (blue) and previously reported (green; monomer C from PDB ID: 3UYS) studies.

Soaking experiments using Epiblastin A (entry **12**, **Table 2**) and derivative **15** (entry **15**, **Table 2**) successfully generated crystals that diffracted at 2.56 Å and 2.76 Å resolution, respectively (see **Figure 29**).

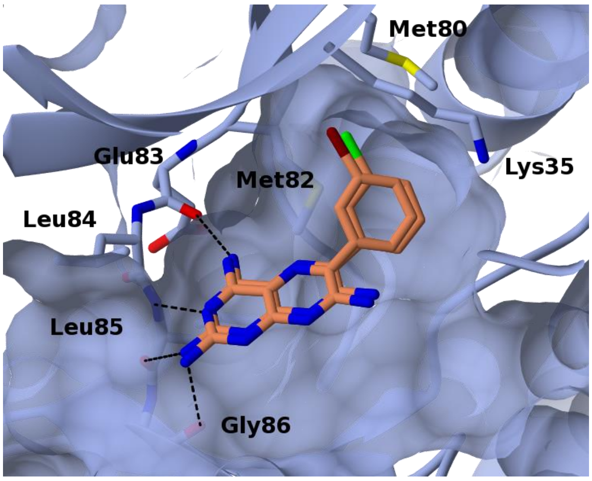


Figure 29. Overlaid crystal structures of Epiblastin A (entry **12**) and derivative **15** (entry **15**) in complex with CSNK1D. Carbon and nitrogen atoms are highlighted in orange and blue, respectively. Chlorine atom in Epiblastin A is shown in green, while the bromine group in derivative **15** is colored brown. Met80, Met82, Glu84, Leu85 and Gly86 point out the amino acid residues that interact with Epiblastin A and derivative **15**, respectively. Lys35 residue is involved in the phosphotransferase activity of CSNK1D.

Both structures show the same crystal forms as previously mentioned for the apo-CSNK1D, namely $P3_121$. The crystal structures of both complexes show a high degree of similarity to the apo-CSNK1D structure as indicated by the root-mean-square deviation (rmsd) (for Epiblastin A: rmsd 0.57 Å over 285 amino acids; for derivative **15**: rmsd 0.40 Å over 276 amino acids). These values suggest that binding of Epiblastin A and derivative **15** does not significantly modify the CSNK1D conformation.

The mode of binding exhibited by Epiblastin A and derivative **15** is perfectly in line with the structure-activity relationship study conclusions generated by the enzymatic study directed towards CSNK1D/E. For example, three atoms from the 2,4-diaminopyrimidine unit form an array of four hydrogen bonds with the hinge region of the CSNK1D nucleotide binding site. In particular, the amino group located at C2 position is involved in two hydrogen bonds with the carbonyl groups of amino acids Leu and Gly located at positions 85 and 86, respectively. The second amino group located at C4 interacts likewise with the carbonyl residue of Glu83. Both Epiblastin A and derivative **15** are additionally docked in the hinge region by means of a fourth hydrogen bond occurring between the nitrogen atom located at position 3 of the pteridine core and the NH backbone of Leu85 that is part of the hinge region. Another particular aspect of the Epiblastin A and derivative **15** mode of binding is the meta-halogen substitution featured by these compounds. The enzymatic studies previously described (see section *4.4. Screening of TR-based derivatives yields potent and / or selective CSNK1D/E inhibitors*), revealed that the presence of the proper atom or group in this position is mandatory to achieve high degree of inhibition of the CSNK1D isoform. Indeed, the crystal structure determinations indicate that the chlorine and bromine atom, respectively are deeply buried in the hydrophobic cavity shaped by the amino acid side chains of Met80 and Met82. Considering the enzymatic studies presented above, the chlorine atom in the meta-position of the phenyl ring at position C6 yielded the most active inhibitor of CSNK1D, followed by the bromine containing derivative. A fluorine atom in the same position generated an almost inactive derivative. These results suggest that the chlorine atom most likely fulfills best the steric and electronic requirements for a tight interaction within the hydrophobic pocket. As shown in **Figure 29**, the presence of the slightly larger bromine atom causes a small conformational change of the pteridine unit in comparison with Epiblastin A. Moreover, a slight shift was observed at the level of the 2,4-diaminopyrimidine unit as well. These aspects could, at least partially, explain the lower potency exhibited by derivative **15** in comparison with Epiblastin A.

Overall, Epiblastin A and derivative **15** (entry **15**) tightly fit into the nucleotide binding pocket as shown in **Figure 30**. This clearly shows that the upper part of the pteridine core encompassing N at position 3, C2, C4 and C6 is directly engaged in non-covalent interactions with the ATP binding pocket of CSNK1D. Therefore, the replacement of these groups will abolish the inhibitory activity in good agreement with the conclusions generated by the enzymatic studies.

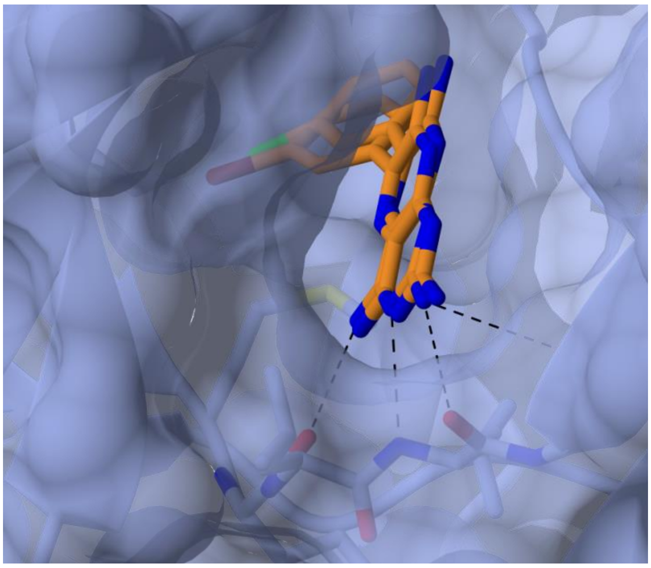
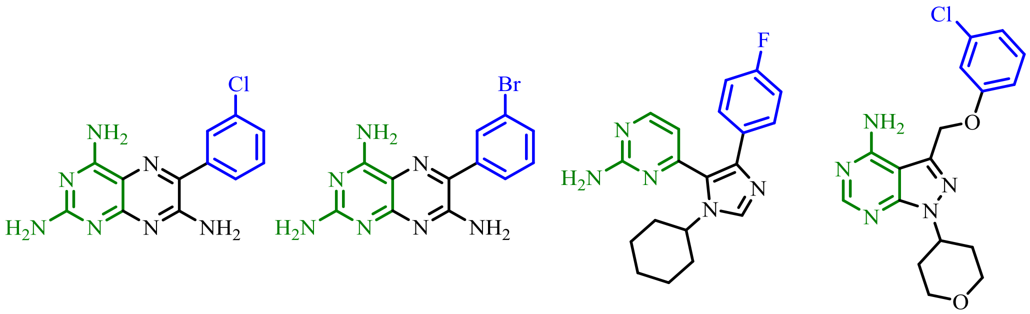


Figure 30. Tight interaction of Epiblastin A and derivative **15** (entry **15**), respectively within the ATP binding site of CSNK1D. Dashed lines indicate the hydrogen-bond network between Epiblastin A and the hinge region.

Having determined the crystal structure in complex with Epiblastin A and derivative **15,** a comparison study with the previously published structures was performed. As a result, the pan-CSNK1D inhibitor PF-670662 (IC_{50} = 13 nM for CSNK1D and IC_{50} = 80 nM for CSNK1E) and the CSNK1E specific inhibitor PF-4800567 (IC_{50} = 711 nM for CSNK1D and IC_{50} = 32 nM for CSNK1E) were selected. Chemical structures for these inhibitors as well as the compounds reported in the current study are shown side by side in **Figure 31**.



Epiblastin A **Derivative 15** **PF-4800567** **PF-670462**

Figure 31. Chemical structures of Epiblastin A, derivative **15**, PF-670662 and PF-4800567. Green color marks the aminopyrimidine unit embedded within the structures of these derivatives. Blue color highlights the halogen substituted phenyl ring of the CSNK1 inhibitors previously mentioned.

At first glance, the CSNK1 inhibitors presented in **Figure 31** share two common features, namely, the aminopyrimidine unit and the halogen substitution on the phenyl ring. A chlorine atom is located in the meta-position of the benzene unit of Epiblastin A and PF-670462, whereas derivative **15** features a bromine atom. However, PF-4800567 exhibits a fluorine-substitution in the para-position of the phenyl ring. As a result of these common features, it is not surprising that the CSNK1 inhibitors previously described share similar binding modes inside the ATP binding pocket, as highlighted in **Figure 32**.

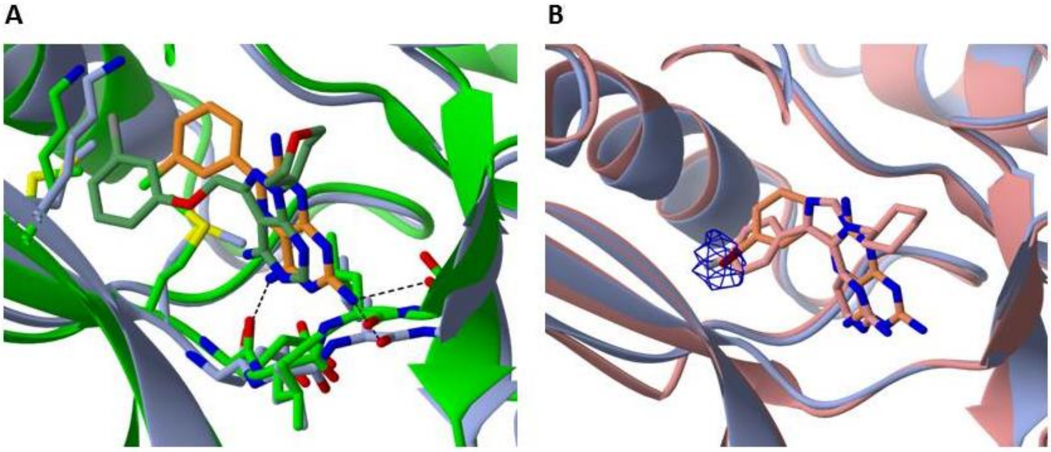


Figure 32. Superimposed crystal structures of Epiblastin A (orange) – PF-4800567 (green) (A) and derivative **15** (orange) – PF-670462 (pink) (B). Carbon and nitrogen atoms are highlighted in orange and blue, respectively. Chlorine atom in Epiblastin A is shown in green, while the bromine group in derivative **15** is coloured brown. The highlighted blue electronic density suggests the anomalous scattering specifically generated by the bromine atom when the wavelength of the X-ray is adjusted to the absorption edge of this atom, namely 0.91883 Å.

As shown by **Figure 32A**, Epiblastin A and PF-4800567 share similar binding modes within the hinge region. However, the meta-chlorine phenyl substituted unit is differently oriented within the hydrophobic pocket. In particular, the chlorine atom points out to opposite directions indicating that the PF-4800567 inhibitor does not engage the hydrophobic pocket created by the Met80 and Met82, as reported above for Epiblastin A and derivative **15**. Conversely, derivative **15** and PF-670462 crystal structures share a higher degree of similarity. As expected, the aminopyrimidine units of both inhibitors are H-bonded with the amino acids within the hinge region. Moreover, the meta-bromo and para-fluoro substitution in the structure of derivative **15** and PF-670462, respectively are oriented towards the hydrophobic pocket shaped by Met80 and Met82 amino acid residues. The anomalous signal highlighted in **Figure 32B** is specifically detected due to the presence of the bromine atom in the structure of derivative **15**. As a result of the strong X-ray absorbance of heavy atoms, anomalous scattering can exclusively be observed when the wavelength of the X-ray matches the absorption edge of the bromine atom, namely 0.91883 Å. This electronic density is located in close proximity to the hydrophobic pocket and suggests a high degree of interaction between derivative **15** with this cavity.

The crystal structure of CSNK1D with PF-670662[217] revealed that the imidazole nitrogen is H-bonded to a water molecule found in close proximity (data not shown). A close inspection of this particular water unit shows that it further interacts with Lys38 and Asp149 amino acid residues that characterize the backbone of the DFG motif. The presence of a water molecule (though with low occupancy) was observed in the vicinity of Epiblastin A and derivative **15** (entry **15**) as well. Unfortunately, it is located too far away for a non-covalent interaction with the amino group at position C7 to occur. On the other hand, this water molecule exhibits the same conserved H-bond network with the amino acid residues of Lys38 and Asp149 previously reported for PF-670662, suggesting that the appropriate modification of the C7 position might be able to engage this water molecule to increase potency and / or selectivity against CSNK1D.

The crystal structures presented herein provide useful insights into the basic requirements for developing potent and / or selective CSNK1D inhibitors. In this regard, further derivatizations at positions C2, C7 and nitrogen atoms at positions 1 and 8 represent the next logical step in exploiting the full potential of the pteridine scaffold.

4.7. Detection of target engagement for Epiblastin A

After profiling the pteridine library *in vitro* against CSNK1A1/D/E isoenzymes, the focus was placed on Epiblastin A. In order to determine whether this compound engages[218] these isoforms in cell lysate, a kinase enrichment experiment (Thermo Scientific) was performed.[219] The protocol makes use of the ActivX™ ATP probe[220,221,222,223] to specifically address ATP binding proteins, generally referred as the purinome.[224,225] This subproteome, represents approximately 13 % of the entire set of cellular proteins[226] and comprises protein kinases, ATPases, helicases, purinergic receptors, motor proteins, various transferases and other proteins.[227] Most of these proteins feature one or two lysine residues located in the close vicinity of the ATP binding pocket.[228,229,230] The ActivX™ ATP probe exploits this feature and contains a labile ester bond located at the γ-phosphate position which terminates in a desthiobiotin moiety connected via a short linker (see chemical structure in **Figure 33**). Similar probes, bearing an alkyne moiety instead of desthiobiotin have also been reported in the literature.[231,232,233]

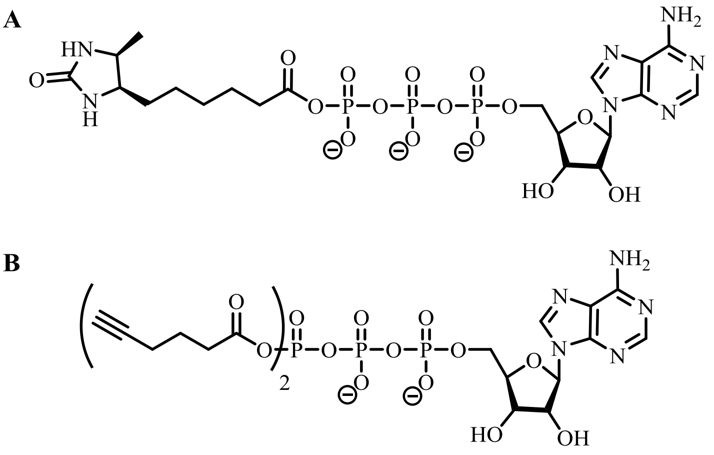


Figure 33. Chemical structures of the ActivX™ ATP probe (A) and ATP-derivative bearing two alkyne units at the γ-phosphate position (B).

Crystal structure determinations of ATP binding proteins elucidated that one or even two lysine residues are located in close proximity to the activated ester. Due to increased effective concentration, which greatly accelerates the reaction rate of labeling, these lysine residues are much more prone to react than, for example, other identical residues located on the protein surface. By combining short incubation times with relatively low concentrations of the ActivX™ ATP probe, the background reactions with other lysine residues can be greatly diminished.[234] After incubation and enrichment using streptavidin coated beads, the labeling efficiency can be evaluated by means of immunoblotting against the target of interest. The ActivX™ ATP probe was previously reported to deliver robust outcomes using different experimental setups, furnishing important insights into inhibitor selectivity[235] and binding constants of small molecules for kinases in cell lysates.[221] These results highlight the versatility and robustness of these chemical probes in revealing novel understandings into how members of the purinome are regulated. The main feature of the ActivX™ ATP probe is the ability to assess the specificity and inhibitory activity of ATP-competitive agents in nucleotide (ATP)-depleted cellular lysates in a similar fashion to competitive activity-based profiling.[236] Therefore, the binding event of a certain compound to an ATP binding protein can be evaluated by means of immunoblotting or proteomic analysis.

Before initiating the target engagement experiments, a suitable cell line had to be identified that contains CSNK1A1/D/E isoforms at detectable levels. Therefore, various cancerous cell lines from different tissue origins such as HeLa, HCT116, OVCAR8, MCF7 and ACHN and the cell line HEK293T were analyzed by means of immunoblotting using specific antibodies. As shown in **Figure 34**, the HCT116 colorectal carcinoma cell line showed a good protein expression profile of CSNK1A1/D/E isoenzymes.

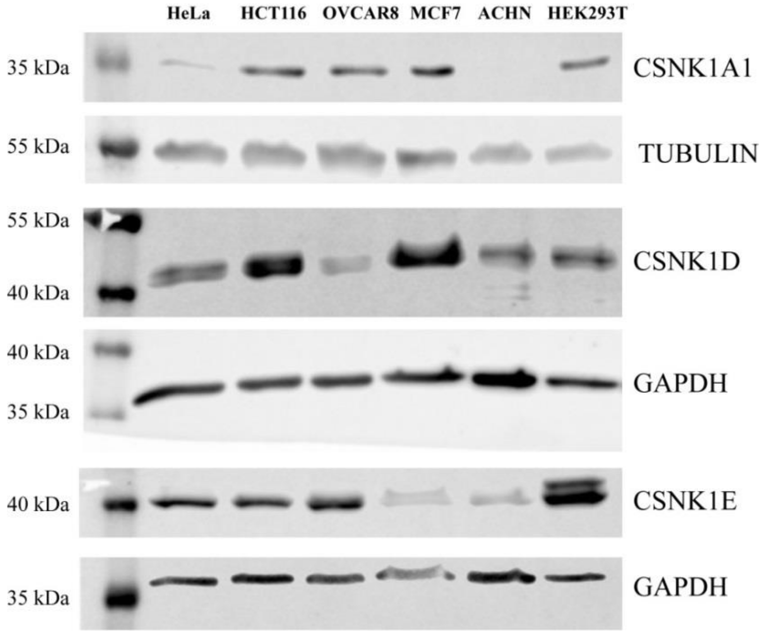


Figure 34. Detection of CSNK1A1/D/E proteins levels in HeLa, HCT116, OVCAR8, MCF7, ACHN and HEK293T cell lines. Proteins were separated by means of SDS-PAGE followed by transfer to a PVDF membrane. CSNK1A1, CSNK1D and CSNK1E were detected using primary anti-CSNK1A1, anti-CSNK1D and anti-CSNK1E antibodies and a secondary antibody coupled to horseradish peroxidase. Tubulin was employed as loading control for CSNK1A1 using anti-α tubulin antibody. GAPDH was acted as loading control for CSNK1D/E using anti-GAPDH antibody. For both controls, the secondary antibody was coupled to horseradish peroxidase.

In addition to the detectable CSNK1A1/D/E protein levels, Epiblastin A altered the proliferation of the HCT116 cell line upon 48 hours incubation as demonstrated by the WST-1 cell viability assay (GI_{50} = 2.46 ± 0.25 μM; GI_{50} is defined as the half-maximal growth inhibitory concentration; see **Figure 35**). This preliminary result indicated the possible crucial involvement of the CSNK1 kinase in HCT116 cell growth (see section *4.11.1. Epiblastin A inhibits the proliferation of wt HCT116 colorectal carcinoma cells*).

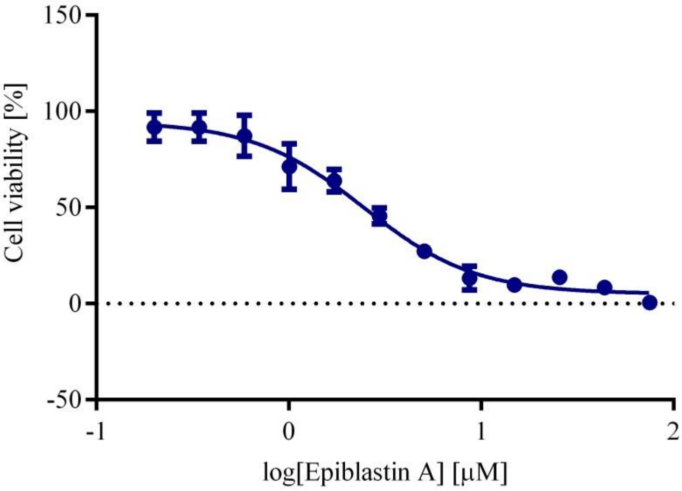


Figure 35. WST-1 cell proliferation assay of wt HCT116 cells upon Epiblastin A treatment for 48 h. Experimental details: 4000 cells were initially plated overnight followed next day by addition of Epiblastin A at various concentrations. Afterwards, WST-1 reagent is added and the absorption at 450 nm is recorded. Cells treated with DMSO were considered 100 % viable. Data represent the average ± SD of three independent experiments performed in quadruplicate. Determined GI_{50} values ± SD were generated by fitting the data with four parameter nonlinear regression analysis using GraphPad Prism version 6.00 for Windows.

As a result of the high CSNK1A1/D/E expression profile at the protein level and the growth inhibition effect upon Epiblastin A treatment, the HCT116 cell line was chosen to test whether Epiblastin A engages CSNK1A1/D/E isoenzymes in cell lysate. Having determined *in vitro* the IC_{50} values for the inhibition of CSNK1A1/D/E isoenzymes, Epiblastin A was employed at 1, 10 and 100 μM, respectively while using 10 μM concentration of the ActivX^{TM} ATP reagent. Therefore, not only the target engagement of a specific inhibitor in cellular lysate could be evaluated in a concentration dependent fashion, but also the dose range at which this compound is effective in cell lysate. The cellular lysates were treated with DMSO (acting as internal control at a final concentration of 1 %) and increasing concentrations of Epiblastin A. Following a short incubation time at room temperature ranging from 10 to 60 min, the ActivX^{TM} ATP probe was added for a defined period of time of either 10 or 60 min. The labeled proteins were then enriched using streptavidin beads, boiled and separated using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Using specific antibodies for the proposed targets, the labeling extent of the DMSO-treated sample can be directly compared with the compound-treated samples. If the inhibitor would compete with the ActivX^{TM} ATP probe for the nucleotide binding site, the intensity of the band revealed by immunoblotting should diminish with increasing concentrations of the inhibitor. Indeed, for Epiblastin A a concentration-dependent protection of CSNK1A1/D/E labeling was observed (see **Figure 36**).

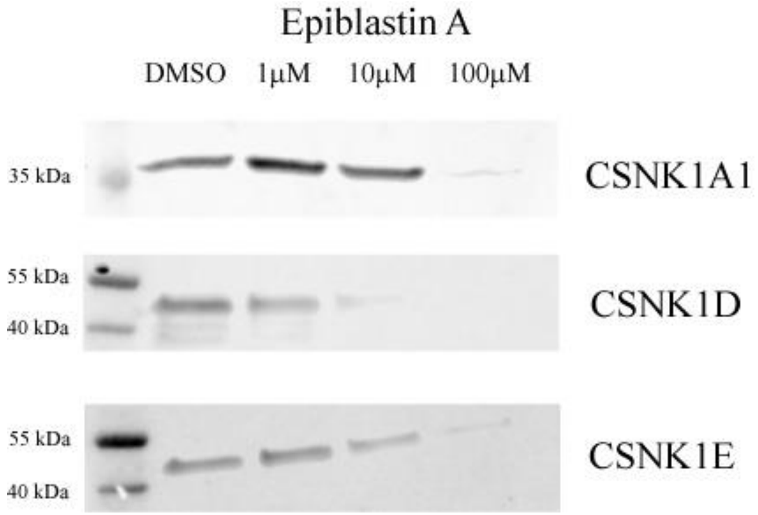


Figure 36. Epiblastin A interferes with the labeling of CSNK1A1/D/E in HCT116 cell lysate. Experimental details: ATP-depleted HCT116 lysates were treated with Epiblastin A (1, 10 and 100 µM) for 10 / 60 min followed by treatment with ActivX^{TM} ATP probe for 10 / 60 min. Labeled proteins were collected using strepatvidin coated beads and separated using SDS-PAGE. Proteins were transferred to a PVDF membrane prior to detection of CSNK1A1, CSNK1D and CSNK1E using anti-CSNK1A1, anti-CSNK1D and anti-CSNK1E primary antibodies and a secondary antibody coupled to horseradish peroxidase.

On one hand, 1 μM and 10 μM of Epiblastin A seemingly did not affect the labeling of CSNK1A1/E in line with the determined IC_{50} values. Thus, it is not surprising that in the context of cell lysate more than 10 μM of the inhibitor is required to detect a noticeable blocking effect. On the other hand, Epiblastin A partially blocked the labeling of CSNK1D at 10 μM compound concentration, in good agreement with the potency *in vitro*. In all cases, adding 100 μM of the compound completely protected all three isoforms from labeling. Overall, these results confirm the engagement of CSNK1A1/D/E family members by Epiblastin A in ATP-depleted cellular lysates and further validate this compound as a novel and selective CSNK1D inhibitor. It is noteworthy to emphasize the fact that these results were generated using ATP-depleted lysates, whereby the nucleotides and other cofactors were partially removed by means of gel filtration.

In conclusion, by means of a competitive activity-based profiling procedure employing the ActivX^{TM} ATP probe and immunoblotting using specific antibodies, the engagement of CSNK1A1/D/E isoenzymes by Epiblastin A in ATP-depleted HCT116 colorectal carcinoma cellular lysates was successfully validated. This method is simple and easy to perform and could be readily applied to confirm the interaction of Epiblastin A with other ATP binding proteins, as discussed in the next section.

4.8. Revealing novel interaction partners of Epiblastin A by means of chemical proteomics

One of the main drawbacks of kinase profiling studies is the limited availability of the naturally occurring enzymes in commercial panels. These kinases are mostly expressed as truncated constructs and purified using various tags, which consequently might alter their native structure and ultimately the enzymatic activity. Encouraged by the results described above that confirmed target engagement of CSNK1A1/D/E members by Epiblastin A using immunoblotting (see previous section), a proteomics-based strategy was applied to unveil novel interaction partners of this compound in HCT116 cell lysates. As compared to the *in vitro* study which mostly analyzes the enzyme activity for a single substrate under defined conditions, the cellular lysate encompasses a more complex environment, whereby multiple proteins simultaneously compete for the binding partners. On one hand, finding additional protein targets for Epiblastin A would be of valuable importance since they might be involved in reprogramming of late-stage mEpiSCs and might explain the enhanced potency of this derivative compared to the structurally similar hit **TR**. On the other hand, these potential targets might elucidate the reduced cell proliferation rate of the HCT116 cells observed upon treatment with Epiblastin A.

However, the strategy based on the ActivX^{TM} ATP probe addresses only a fraction of the proteome, namely the purinome. As a consequence, other target classes are entirely omitted from the analysis. Conversely, focusing on a rather small fraction of the proteome would be advantageous, since highly expressed background contaminants could be largely removed from this procedure. In this way, even low expressed proteins, such as kinases, could be readily isolated and subsequently identified. Due to the crucial role of kinases in various aspects of stem cell biology and cancer research, the purinome-focused strategy might suit the current research interests. Therefore, after labeling with the ActivX^{TM} ATP based probe, enrichment using streptavidin coated beads and boiling with Laemmli buffer, the proteins were separated by means of SDS-PAGE and digested with trypsin using an in-gel digestion procedure. The digested peptides were then analyzed using nano-LC / MS followed by label-free quantification (LFQ).

This protocol was applied under four slightly different conditions (see below). Since 100 μM Epiblastin A produced a noticeable effect in the immunoblotting experiments, this concentration was considered sufficient to detect changes in the peptide intensities generated during the MS procedure compared to samples treated with DMSO. Thus, the concentrations of Epiblastin A and the ActivX^{TM} ATP probe were fixed at 100 μM and 10 μM, respectively, coupled with varying incubation times of these two components.

Protein code	Protein name	10 min / 10 min	10 min / 60 min	60 min / 10 min	60 min / 60 min
STK10	Serine / threonine-protein kinase 10	XX	X	XX	XX
CSNK1A1	Casein kinase I isoform alpha	X	XX	X	X
PIP4K2A	Phosphatidylinositol 5-phosphate 4-kinase type-2 alpha	XX	X	XX	X
PIP4K2C	Phosphatidylinositol 5-phosphate 4-kinase type-2 gamma		X	X	X
PRKDC	DNA-dependent protein kinase catalytic subunit	XX			
DYNC1H1	Cytoplasmic dynein 1 heavy chain 1	XX			
PRKACB	cAMP-dependent protein kinase catalytic subunit beta		XX		
ATP6V0D1	V-type proton ATPase subunit d 1		XX		
MAPK8	Mitogen-activated protein kinase 8		X		
MVD	Diphosphomevalonate decarboxylase		X		
MAP2K4	Dual specificity mitogen-activated protein kinase kinase 4			X	
GNE	Bifunctional UDP-N-acetylglucosamine 2-epimerase/ N-acetylmannosamine kinase	X			
EHD4	EH domain-containing protein 4	X			
TTL	Tubulin-tyrosine ligase				X

Table 8. Proteins identified to interact with 100 μM Epiblastin A using 10 μM of the ActivX^{TM} ATP probe at specific time points in HCT116 cell lysate. The proteins marked with X were found significant by the median test statistical analysis, while the targets labeled with XX were very significant under the same statistical parameters; x min / x min suggests the incubation time of the HCT116 cell lysate with Epiblastin A and ActivX^{TM} ATP probe, respectively.

The experimental conditions imply the incubation with Epiblastin A for 10 and 60 minutes respectively at 100 μM followed by the treatment with ActivX^{TM} ATP probe at 10 μM for 10 and 60 minutes respectively. As a consequence, four experiments (designated in **Table 8** as 10 / 10, 10 / 60, 60 / 10 and 60 / 60, respectively) were performed that would offer not only a better coverage of the possible targets but also provide snapshots of the Epiblastin A interactomes (protein targets that interact with Epiblastin A) in cellular lysate in a time-dependent fashion. This in turn could uncover which targets are modulated in the first 10 minutes of incubation and possible differences in the target profile after 60 minutes of Epiblastin A treatment in cellular lysates. **Table 8** shows the proteins that were identified to bind to Epiblastin A under the given experimental conditions. A complete list of the identified ATP-binding and non-ATP binding proteins can be found in **Table 11** (see section *8. Appendix).*

Subjection of the protein list shown in **Table 8** to the CRAPome (Contaminant Repository for Affinity Purification)[237] database that contains lists of common background contaminants identified in affinity-purification mass-spectrometry (AP-MS) experiments. With the exception of two proteins, namely DYNC1H1 (Cytoplasmic dynein 1 heavy chain 1) and PRKDC (DNA-dependent protein kinase catalytic subunit, also known as DNA-PK) that exhibited relatively high expectancy, the rest of the identified proteins was not predicted as background contaminants (data not shown). Taking in consideration the nature of the experiment, which implies stringent washing steps with 4 M urea, and the fact that DYNC1H1 and PRKDC only appeared once in these experiments with high statistical coverage, suggests that they might indeed be genuine protein targets of Epiblastin A. Moreover, to find out if the list contains proteins that might be co-enriched due to the formation of macromolecular protein complexes, the same list was also submitted to the String database,[238] which encompasses known and predicted protein-protein interactions. No direct association could be found between the targets in single experiments, suggesting the validity of the protein list. Two exceptions were identified as PRKDC-MAPK8 (Mitogen-activated protein kinase 8) and MAPK8-MAP2K4 (Dual specificity mitogen-activated protein kinase kinase 4), but these proteins were identified in different experiments. Since PRKDC and MAPK8 were previously associated with the modulation of DNA damage and stress signaling, respectively, one can only speculate about their inhibition upon Epiblastin A treatment as a possible mechanism of action associated with this compound.

At a first glance **Table 8** highlights an important feature of the obtained results. There appear to be two types of targets addressed by Epiblastin A. On one hand, STK10 (Serine / threonine-protein kinase 10), CSNK1A1 (Casein kinase I isoform alpha), PIP4K2A (Phosphatidylinositol 5-phosphate 4-kinase type-2 alpha) and PIP4K2C (Phosphatidylinositol 5-phosphate 4-kinase type-2 gamma) are significantly present in each one of the experiments performed, which might suggest a high degree of interaction regardless of the incubation times with Epiblastin A and the ActivX^{TM} ATP probe. A second set of targets was identified only once indicating transient association with Epiblastin A, solely at specific time points.

Since all the experimental conditions are identical excepting the incubation times, it can be argued that the protein list in **Table 8** depicts a time-dependent snapshot of the Epiblastin A interactome in HCT116 cell lysate. Furthermore, since the cellular lysates represent just an averaged protein concentration across a cell population, this scenario might dramatically differ from the context of a living cell, where the system is highly dynamic and protein levels are modulated in a spatio-temporal fashion.

Of the main targets CSNK1A1, STK10, PIP4K2A and PIP4K2C, only the CSNK1A1 could be identified in each of the four experiments, confirming the target engagement results described above using the immunoblotting procedure. However, the chemical proteomics experiments could not detect either of the CSNK1D or CSNK1E isoenzymes. This does not imply that the interaction between the two isoforms and Epiblastin A does not occurr in the cellular context, but rather stresses the lower abundance of these proteins that might not be detectable by the mass spectrometer under the experimental conditions. Thus, judging by the chemical proteomics results, CSNK1A1 is the main target addressed by Epiblastin A in HCT116 cell lysates among the CSNK1 kinase family.

STK10 was identified as a prominent target during the study, considering the sequence coverage and signal intensities. STK10, a polo-like kinase kinase,[239] represents the human homologue of the murine Lymphocyte oriented kinase (LOK) and is part of the Ste20 serine / threonine kinase family. To confirm STK10 as a *bona fide* target of Epiblastin A, the inhibitory activity of Epiblastin A on STK10 was investigated (see **Figure 37**).

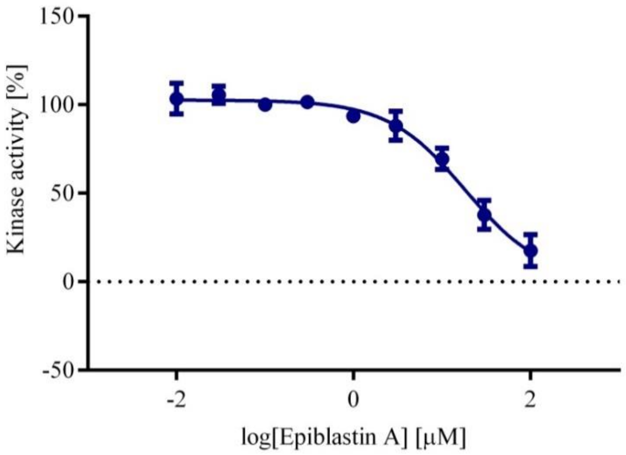


Figure 37. IC_{50} determination of Epiblastin A for inhibition of STK10 (h) determined by mean of radiometric-based assay at K_m-ATP value. Data represent mean values ± SD of three independent measurements. Determined IC_{50} value ± SD were generated by fitting the data with three parameter nonlinear regression analysis using GraphPad Prism version 6.00 for Windows.

Concentration-dependent measurement performed at Eurofins (former Milipore) indicated an IC_{50} value of 18 μM, confirming STK10 as a novel target addressed by Epiblastin A *in vitro* Additionally, the interaction of Epiblastin A with STK10 in HCT116 lysates was demonstrated by means of competition with the ActivX^{TM} ATP followed by immunoblotting using the procedure described above for the CSNK1A1/D/E isoenzymes (see **Figure 38**).

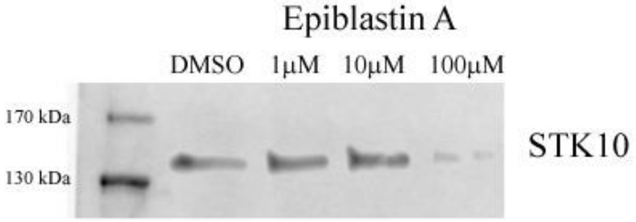


Figure 38. Epiblastin A disturbs the labeling of STK10 in HCT116 cell lysates. Experimental details: nucleotide-depleted lysates of HCT116 cells were incubated with Epiblastin A (1, 10 and 100 μM) for 10 / 60 min followed by treatment with the ActivX^{TM} ATP probe for 10 / 60 min. Labeled proteins were enriched using streptavidin beads that were separated by means of SDS-PAGE. Proteins were transferred to a PVDF membrane prior to detection of STK10 using anti-STK10 antibody and a secondary antibody coupled to horseradish peroxidase.

Not surprisingly, Epiblastin A exhibited a similar inhibition behavior for the labeling of STK10, namely more than 10 μM of this compound is necessary to observe a noticeable protection from labeling with the ActivX^{TM} ATP probe. This further validates that the developed chemical proteomics method can reveal novel target proteins of Epiblastin A.

4.9. Data analysis of chemical proteomics experiments

The four chemical proteomics experiments identified between 1187 and 1812, proteins with an average of 1440 proteins per experiment (see **Table 9**).

	10 / 10	**10 / 60**	**60 / 10**	**60 / 60**
Total number of proteins	1187	1383	1378	1812
ATP binding proteins	327	349	338	408
Kinase activity	154	168	162	198
ATP binding and kinase activity	125	134	131	160
GTP binding	61	75	78	93

Table 9. Number and type of proteins identified during the chemical proteomics experiments. Proteins identified to fulfill a specific biological role were selected by means of the molecular function gene ontology (GOMF) classification. x / x suggests the incubation time of the HCT116 cell lysate with Epiblastin A and the ActivX^{TM} ATP probe, respectively.

From the total amount of proteins, an average of 355 was designated as ATP-binding proteins by the molecular function gene ontology (GOMF) analysis. Thus, more than 25 % of the total amount of proteins identified is reported to bind ATP. As a comparison, only 77 GTP-binding proteins could be detected representing less than 5 % of the all identified proteins suggesting a high degree of enrichment of ATP-binding proteins. In addition, approximately 40 % of the identified targets that interact with ATP are reported to possess kinase activity as mentioned by the GOMF terms which corresponds to roughly137 kinases per experiment. Taking in consideration that the human genome encompasses more than 500 kinases, the chemical proteomics experiments based on ActivX^{TM} ATP probe was able to enrich approximately 27 % of the whole kinome from 1 mg of HCT116 protein lysate. As previously reported in the literature,[221] increasing the amount of cell lysate to 5 mg together with more sensitive MS detection can increase the overall coverage of kinases, thus allowing for an improved profiling capability of a small molecule across the entire kinome, for example.

In conclusion, the ActivX^{TM} ATP probe represents a robust chemical reagent that allows approximately 25 % enrichment of the purinome across the identified proteins. Approximately 40 % of the entire ATP binding proteins isolated during the chemical proteomics procedures proved to possess kinase activity. Thus, the chemical proteomics procedure based on ActivX^{TM} ATP probe offers a suitable method to profile ATP-competitive inhibitors across the purinome and to detect target engagement of these derivatives in ATP-depleted cellular lysates.

4.10. Kinase profiling reveals targets of Epiblastin A potentially involved in reprogramming

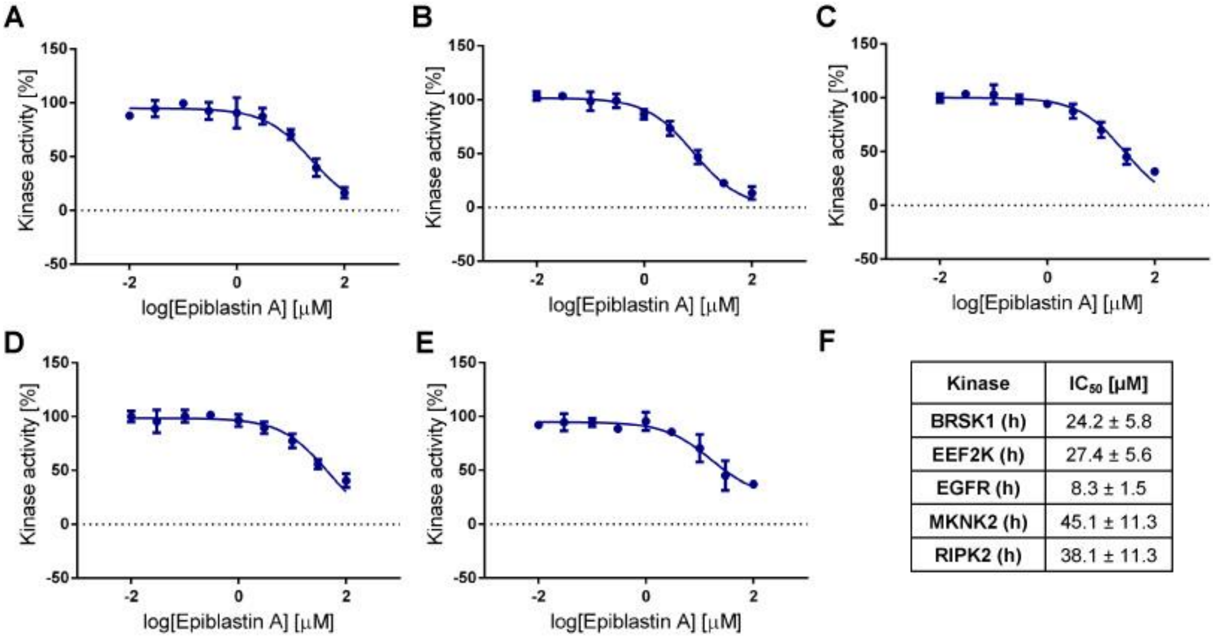
As shown above, the stem cell reprogramming activity of **TR** and Epiblastin A is due to their ability to inhibit CSNK1A1/D/E with high specificity.[148,149] In particular, Epiblastin A is a highly specific CSNK1A1/D/E inhibitor in a panel comprised of more than 100 kinases including members of the major signaling pathways reported to affect stem cell fate.[148,149] Moreover, both compounds also targeted the PI3K family members, albeit to different extents. In order to determine whether the compounds might inhibit other kinases that were not identified in the first screening campaign or during the chemical proteomic experiments, Epiblastin A's inhibitory activity was investigated in single-point measurements at 10 μM and K_m-ATP value for 123 additional kinases. The rationale in selecting these targets was to cover as many kinase families as possible. Moreover, if several members within the same subfamily were available in the panel, only the most representative examples were selected that exhibit different sequences within the ATP-binding pocket region. The assay was performed at Eurofins (former Millipore) using a ^{33}P-radiometric-based assay. This study revealed more than 50 % inhibition of brain-selective kinase 1 (BRSK1), eukaryotic elongation factor 2 kinase (EEF2K), epidermal growth factor receptor kinase (EGFR), MAP kinase-interacting serine / threonine-protein kinase 2 (MKNK2) and receptor-interacting serine / threonine-protein kinase 2 (RIPK2) (see **Table 10**).

Kinase	**Residual activity (%)**	**SD**	**Kinase**	**Residual activity (%)**	**SD**
ACK1(h)	80	0	Mer(h)	84	1
Arg(h)	115	7	Met(h)	94	5
AMPKα1(h)	109	5	MINK(h)	95	10
AMPKα2(h)	123	10	MLCK(h)	78	7
ASK1(h)	90	1	Mnk2(h)*	56	2
Axl(h)	93	3	MRCKα(h)	86	5
Bmx(h)	92	3	MSK1(h)	89	1
BRK(h)	103	9	MSK2(h)	102	7
BrSK1(h)*	64	3	MSSK1(h)	105	6
BTK(h)	123	17	MST4(h)	90	3
CaMKI(h)	98	11	MuSK(h)	91	3
CaMKIIβ(h)	101	5	NEK9(h)	98	6
CaMKIIδ(h)	99	3	PAK1(h)	88	8
CaMKIV(h)	106	1	PAR-1Bα(h)	104	1
CHK1(h)	99	6	PASK(h)	105	8
CHK2(h)	112	8	PEK(h)	96	3
cKit(h)	102	8	PhKγ2(h)	99	2
CSK(h)	105	2	Pim-1(h)	95	4
DAPK1(h)	83	4	Pim-2(h)	95	2
DAPK2(h)	116	11	Pim-3(h)	82	7
DCAMKL2(h)	110	3	PKD2(h)	83	3
DDR2(h)	103	16	PKG1α(h)	87	9
DMPK(h)	103	9	PKG1β(h)	96	6
DRAK1(h)	86	4	Plk1(h)	95	2
DYRK2(h)	87	2	Plk3(h)	92	2
eEF-2K(h)*	56	8	PRAK(h)	88	10
EGFR(h)*	33	1	PRK2(h)	78	2
EphA1(h)	100	1	PrKX(h)	85	3

Kinase	**Residual activity (%)**	**SD**	**Kinase**	**Residual activity (%)**	**SD**
EphA2(h)	87	1	Pyk2(h)	112	7
EphA3(h)	117	11	Ret(h)	74	2
EphA4(h)	100	2	RIPK2(h)*	49	1
EphA5(h)	95	2	Ron(h)	88	5
EphA7(h)	96	5	Rse(h)	93	12
EphA8(h)	103	7	Rsk2(h)	97	8
EphB1(h)	89	10	Rsk3(h)	94	2
ErbB4(h)	101	1	SGK2(h)	110	8
FAK(h)	88	15	SGK3(h)	86	2
Fes(h)	95	3	SIK(h)	77	5
Flt1(h)	100	6	Snk(h)	91	1
Flt3(h)	102	6	SNRK(h)	81	6
Flt4(h)	84	1	Src(1-530)(h)	87	1
Fms(h)	106	0	LOK(h)*	55	10
Fyn(h)	103	12	LRRK2(h)	77	7
GCK(h)	100	1	MAPKAP-K3(h)	94	13
GCN2(h)	95	4	MARK1(h)	94	13
GRK1(h)	85	5	SRPK1(h)	101	2
GRK6(h)	94	3	STK25(h)	96	0
Hck(h)	81	5	Tec(h) activated	90	0
HIPK1(h)	103	4	TGFBR1(h)	116	4
IGF-1R(h)	95	5	Tie2 (h)	78	1
IGF-1R(h), activated	95	3	TLK1(h)	108	2
IKKε(h)	97	5	TrkA(h)	99	15
IR(h)	83	8	Txk(h)	90	5
IR(h), activated	101	6	TYK2(h)	82	11
IRE1(h)	93	9	ULK2(h)	85	3
IRR(h)	98	4	ULK3(h)	98	6
Itk(h)	96	0	Wee1(h)	85	9
JAK1(h)	100	3	WNK2(h)	91	3
Lck(h)	81	9	WNK3(h)	86	5
Lck(h) activated	115	12	ZAP-70(h)	93	6
LKB1(h)	92	1	ZIPK(h)	111	10

Table 10. Influence of 10 μM Epiblastin A on the activity of 123 selected kinases. Kinase activities were determined at Eurofins (former Milipore) by means of a radiometric-based assay using K_m-ATP for each enzyme. Data are mean values of two independent measurements. Kinases marked with * showed 50 % enzyme inhibition upon treatment with 10 μM Epiblastin A and were further selected for concentration dependent studies (see **Figure 39**).

Analysis of these kinases in concentration-depended manner for inhibition by Epiblastin A under the same experimental conditions are described below (see **Figure 39**).



Kinase	IC_{50} [μM]
BRSK1 (h)	24.2 ± 5.8
EEF2K (h)	27.4 ± 5.6
EGFR (h)	8.3 ± 1.5
MKNK2 (h)	45.1 ± 11.3
RIPK2 (h)	38.1 ± 11.3

Figure 39. IC_{50} determinations for Epiblastin A against BRSK1 (A), EEF-2K (B), EGFR (C), MKNK2 (D) and RIPK2 (E), respectively evaluated at K_m-ATP of each kinase. Data are mean values of three independent measurements performed at Eurofins by mean of radiometric-based assay. (F) Determined IC_{50} values ± SD were generated by fitting the data with three parameter nonlinear regression analysis using GraphPad Prism version 6.00 for Windows.

With the exception of EGFR (IC_{50} = 8.3 ± 1.5 μM), Epiblastin A exhibited higher half-maximal inhibitory concentration values for all other protein kinases compared to the CSNK1 family (see **Figure 39**). This finding might suggest that BRSK1, EEF2K, MKNK2 and RIPK2 inhibition might not be significantly involved in the reprogramming mechanism, although target engagement experiments in mouse epiblast stem cell lysates are required to completely validate these aspects.

4.11. Investigation of Epiblastin A in cancer cell lines

4.11.1. Epiblastin A inhibits the proliferation of wt HCT116 colorectal carcinoma cells

As a result of various studies connecting CSNK1 and cancer research,[199] the effect of Epiblastin A on the proliferation (in this subchapter the terms proliferation and growth are interchangeable) of wt HCT116 (wt refers to HCT116 parental cell line that contains both

TP53 alleles, i.e. *TP53*$^{+/+}$) colorectal carcinoma cells was investigated in-depth. In this particular cell line, CSNK1A1/D/E could be detected at the protein level as previously shown (see **Figure 36**, section *4.7. Detection of target engagement for Epiblastin A*). Moreover, preliminary studies have shown that Epiblastin A inhibited the proliferation of wt HCT116 cells upon 48 h treatment (GI_{50} = 2.46 ± 0.25 μM, GI_{50} is defined as the half-maximum inhibitory cell growth concentration) as indicated by the WST-1 cell proliferation assay (see **Figure 35**, section *4.7. Detection of target engagement for Epiblastin A*).

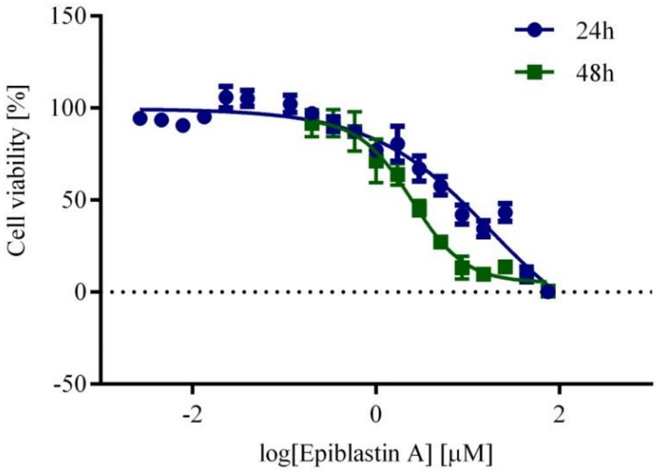


Figure 40. WST-1 cell proliferation assay of wt HCT116 cells upon Epiblastin A treatment for 24 h and 48 h incubation, respectively. Experimental details: 4000 cells were initially plated overnight followed next day by addition of Epiblastin A at various concentrations. After overnight incubation, WST-1 reagent is added and the absorption at 450 nm is recorded. Cells treated with DMSO were considered 100 % viable. Data represent the average ± SD of three independent experiments performed in quadruplicate. Determined IC_{50} values ± SD were generated by fitting the data with four parameter nonlinear regression analysis using GraphPad Prism version 6.00 for Windows.

On the other hand, incubation of wt HCT116 cells with Epiblastin A for 24 h yielded a significantly higher half-maximum inhibitory growth inhibition value (GI_{50} = 18.28 ± 8.63 μM), suggesting that prolonged treatment with this derivative is required to markedly affect the proliferation of wt HCT116 cells (see **Figure 40**).

WST-1 reagent evaluates the number of metabolically active cells, i.e. viable cells in cell culture by measuring the enzymatic activity of mitochondrial dehydrogenases. Although the assay is relatively robust, it cannot exactly explain the cause of the reduced WST-1 signal arising upon treatment with a cytotoxic agent. For example, a decrease in the intensity can be caused by either a direct inhibition of the mitochondrial dehydrogenases, an inhibition of other cellular components that directly regulate the mitochondrial dehydrogenases levels or due to a reduced number of cells in culture. In order to verify the latter hypothesis, this number can easily be approximated using Tryphan blue. This diazo dye specifically stains dead cells. Viable cells however do not allow the dye to pass the cellular membrane and thus are not labeled. As a result, this dye exclusion method allows the possibility to directly asses the number of dead cells in cell culture upon treatment with small molecules (see **Figure 41**).

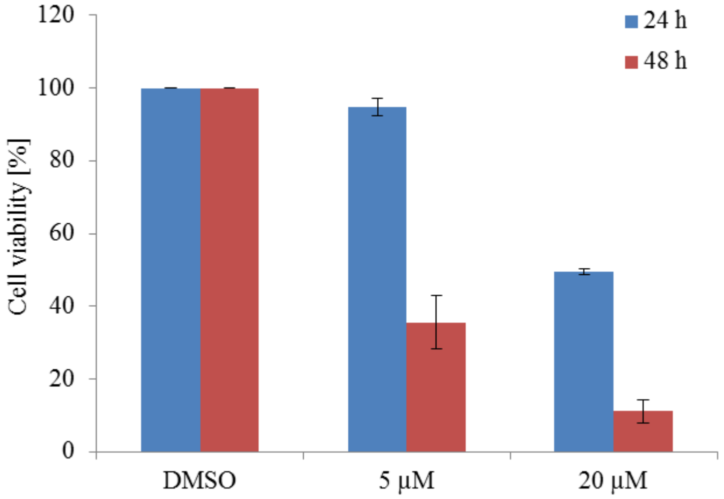


Figure 41. Cell viability evaluated by means of Tryphan blue upon treating wt HCT116 cells with 5 µM and 20 µM Epiblastin A for 24 h and 48 h, respectively. Experimental details: 4000 cells were initially plated overnight followed next day by addition of Epiblastin A at 5 µM and 20 µM for 24 h and 48 h, respectively. After incubation, cells are tryspinized and treated with Tryphan blue. The cells are counted with Countess automated cell counter (Life Technologies). Cells treated with DMSO control were considered 100 % viable. The results represent the average ± SD of three independent experiments performed in duplicate.

The results of the Tryphan blue experiment pointed out the following observations: (1) 5 µM Epiblastin A does not significantly reduce the number of viable wt HCT116 cells upon 24 h treatment, whereas 20 µM reduce the total number to almost half; (2) 5 µM Epiblastin A significantly reduces the number of viable wt HCT116 cells only upon 48 h treatment, whereas higher concentrations, namely 20 µM almost completely blocked cell growth. These conclusions are in very good agreement with the GI_{50} values obtained from WST-1 assay and indicate a direct correlation between the mitochondrial dehydrogenase enzymatic activity measured by this assay and number of viable cells evaluated by means of Tryphan blue.

Next, the effect of Epiblastin A treatment on the cellular ATP content was directly measured in wt HCT116 culture. For this, CellTiter-Glo®,[240] a commercial available cell viability assay that uses luminescence as readout, was employed. Compared to WST-1, the CellTiter-Glo® reagent quantifies the number of viable cells by means of the cellular ATP content. Upon adding the reagent in the cell culture, cell lysis occurs and the released ATP, in the presence of luciferin and molecular oxygen, is converted to AMP and oxyluciferin. As a result of this reaction, a "glow-type" luminescent signal is generated. Thus, the intensity of the luminescent signal is directly proportional to the cellular ATP content in culture. Dose-dependent treatment of wt HCT116 cells with Epiblastin A for 24 h and 48 h, respectively, yielded the curves as shown in **Figure 42**.

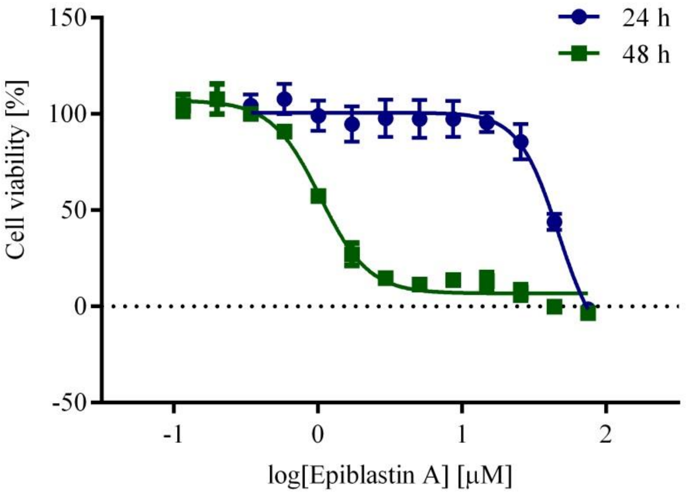


Figure 42. Cell viability evaluated by means of CellTiter-Glo® reagent upon treating wt HCT116 cells with Epiblastin A for 24 h and 48 h, respectively. Experimental details: 4000 cells were initially plated overnight followed next day by addition of Epiblastin A at various concentrations for 24 h and 48 h, respectively. Then, CellTiter-Glo® reagent was added and luminescence was recorded. Cells treated with DMSO were considered 100 % viable. Data represent the average ± SD of three independent experiments performed in duplicate. Determined IC_{50} values ± SD were generated by fitting the data with four parameter nonlinear regression analysis using GraphPad Prism version 6.00 for Windows.

The CellTiter-Glo® experiment suggested that ATP content across the entire cell population is not significantly affected upon Epiblastin A treatment for 24 h (GI_{50} = 46.61 ± 7.74 μM). Only concentrations higher than 50 μM caused a dramatic reduction of the metabolically active cells. On the other hand, 48 h of Epiblastin A treatment significantly diminished the ATP levels in cell culture (GI_{50} = 1.03 ± 0.06 μM) to a similar extent as shown for WST-1 and Tryphan blue experiments previously discussed. Thus, treating wt HCT116 cells with Epiblastin A causes a decreased activity of the mitochondrial dehydrogenases and an overall diminished ATP content across the cell population. These outcomes might be due to a reduced cell number compared to the DMSO treated sample.

CellTox™ Green cytotoxicity assay was further employed to investigate the mode of action of Epiblastin A. This assay contains an asymmetric cyanine dye that detects changes in the membrane integrity of cells directly in culture. If treatment with a small molecule results in cell death, the cellular membrane disintegrates and thus the dye can easily enter the cell and bind to DNA. As a consequence of the binding process, the fluorescent properties of the dye are turned on. Therefore, the fluorescent signal produced is proportional to the number of dead cells across the entire population. Conversely, the dye is completely excluded from viable cells that possess an intact cell membrane which does not allow the dye to pass through. As a consequence, the cyanine dye cannot bind to the DNA anymore and therefore no fluorescent signal can be detected. Epiblastin A was incubated with wt HCT116 cells for 24 h and 48 h, respectively and the results of the CellTox^{TM} Green assay are presented in **Figure 43**. The CellTox^{TM} Green experiments nicely parallels the results obtained with WST-1 and CellTiter-Glo$^{®}$ viability assays as previously discussed.

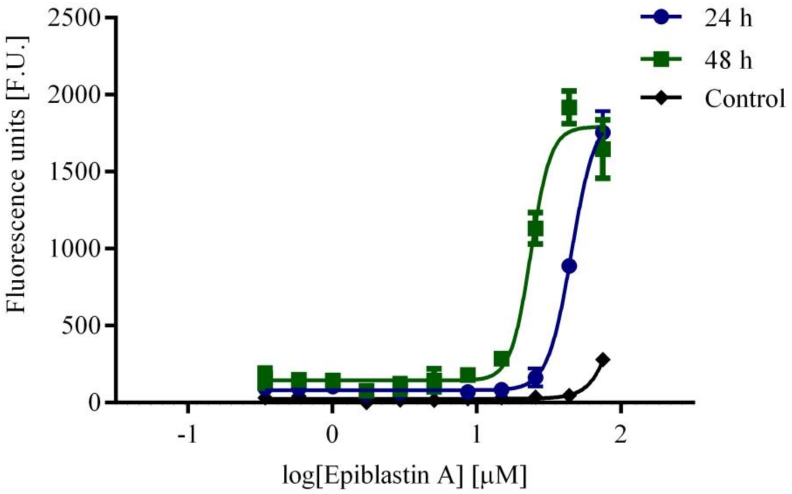


Figure 43. Cell toxicity evaluated by means of CellTox^{TM} Green reagent upon treating wt HCT116 cells with Epiblastin A for 24 h and 48 h, respectively. CellTox^{TM} Green reagent was added and fluorescence emission was recorded at 520 nm after exciting the sample with 485 nm light. Cells treated with DMSO were considered 100 % viable. Data represent the average ± SD of three independent experiments performed in duplicate. Determined IC_{50} values ± SD were generated by fitting the data with four parameter nonlinear regression analysis using GraphPad Prism version 6.00 for Windows.

Apparently, cell death completely occurs after 24 h only at high Epiblastin A concentrations such as 75 μM. In the same time, 50 μM of Epiblastin A induced cell death to approximately half of the wt HCT116 cell population (GI_{50} = 45.43 ± 1.75 μM). Even after 48 h of incubation, significant membrane disintegration characteristic to cells undergoing cell death can only be observed for Epiblastin A at concentrations higher than 25 μM (GI_{50} = 23.70 ± 1.00 μM). Finally, to test if the growth inhibition mechanism is mediated by CSNK1 kinase family, the pan-CSNK1 inhibitor D4476[200] was tested in the WST-1 assay (see **Figure 44**).

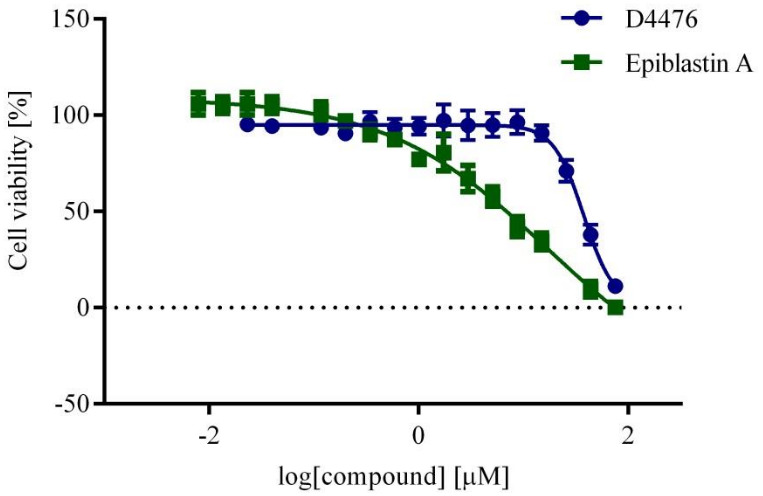


Figure 44. WST-1 cell proliferation assay of wt HCT116 cells upon Epiblastin A and D4476 treatment for 24 h. Experimental details: 4000 cells were initially plated overnight followed next day by addition of Epiblastin A or D4476 at various concentrations. After overnight incubation, WST-1 reagent is added and the absorption at 450 nm is recorded. Cells treated with DMSO were considered 100 % viable. Data represent the average ± SD of three independent experiments performed in quadruplicate. Determined IC_{50} values ± SD were generated by fitting the data with four parameter nonlinear regression analysis using GraphPad Prism version 6.00 for Windows.

The dose-response curve yielded a GI_{50} value of 37 μM (GI_{50} = 37.34 ± 2.81 μM) for this derivative. Under the same experimental conditions, Epiblastin A yielded a half-inhibitory cell growth concentration of 17 μM (GI_{50} = 16.89 ± 7.00 μM). This result preliminarily indicates that the mechanism of cell growth inhibition mediated by Epiblastin A might not be due to the inhibition of the enzymatic activity of CSNK1 kinase family. Further experiments are required to fully investigate this aspect.

In conclusion, Epiblastin A inhibits the cell proliferation of the colorectal carcinoma wt HCT116 cell line. As indicated by the WST-1 assay, Epiblastin A treatment only slightly reduces the activity of the mitochondrial dehydrogenases upon 24 h treatment. Similarly, the ATP content does not seem to be affected at low Epiblastin A concentrations as suggested by the CellTiter-Glo® assay. Only at concentrations higher than 50 μM significant reduction in the cellular ATP content across the wt HCT116 cell population can be observed. Additionally, significant membrane disruptions characteristic to cells undergoing cellular death occurred only with high concentrations of Epiblastin A such as 50 μM as showed by the CellTox™ Green assay. However, treating wt HCT116 cells with 25 μM Epiblastin A for 48 hours induced significant lost in mitochondrial dehydrogenase activity and ATP content that subsequently lead to membrane disintegration characteristic to dead cells. Further assays are needed to confirm if Epiblastin's A mode of action involves programmed cell death, namely apoptosis.

4.11.2. Growth inhibition of HCT116 cells by Epiblastin A is influenced by various parameters

Having determined that Epiblastin A inhibits the proliferation of wt HCT116 colorectal carcinoma cell line, the influence of various parameters was investigated. Factors such as cell number and *TP53* gene status were analyzed as shown in the following paragraphs.

To investigate the effect of the cell number on the growth inhibition mechanism promoted by Epiblastin A, various wt HCT116 cell numbers ranging from 4,000 to 20,000 cells were seeded overnight. Then, Epiblastin A was added for 24 h followed by addition of the WST-1 reagent (see **Figure 45**). Apparently, changing the number of the initially plated cells positively influences the cell growth inhibition mechanism observed upon Epiblastin A treatment.

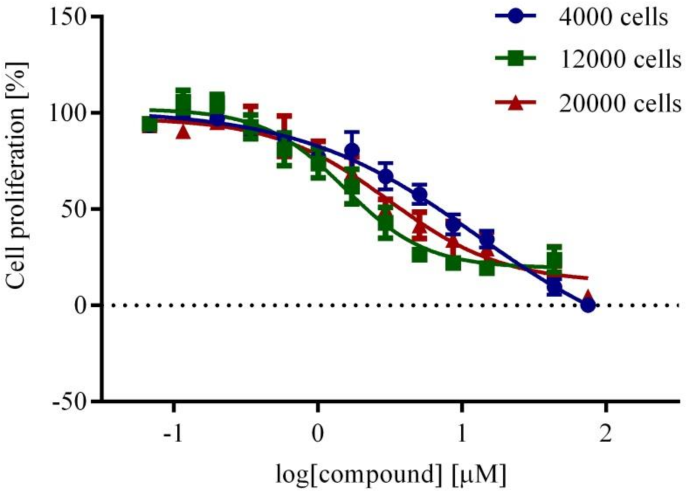


Figure 45. WST-1 cell proliferation assay of wt HCT116 cells upon Epiblastin A treatment for 24 h. Experimental details: 4,000, 12,000 and 20,000 cells respectively were initially plated overnight followed next day by addition of Epiblastin A at various concentrations. After overnight incubation, WST-1 reagent is added and the absorption at 450 nm is recorded. Cells treated with DMSO were considered 100 % viable. Data represent the average ± SD of three independent experiments performed in quadruplicate. Determined IC_{50} values ± SD were generated by fitting the data with four parameter nonlinear regression analysis using GraphPad Prism version 6.00 for Windows.

Unexpectedly, the increase from 4,000 to 12,000 cells significantly reduced the GI_{50} value (from GI_{50} = 12.91 ± 4.29 µM for 4,000 cells to GI_{50} = 1.55 ± 0.18 µM for 12,000 cells). Plating 20,000 cells did not dramatically modify this value suggesting a saturation effect (GI_{50} = 1.55 ± 0.18 µM for 12,000 cells to GI_{50} = 3.12 ± 0.49 µM for 20,000 cells).

In conclusion, the different cell density of wt HCT116 cells initially plated does interfere with the mode of action of Epiblastin A. Further investigations in this direction shall focus on the involvement of cell-cell contacts in enhancing the cell proliferation blocking effect upon Epiblastin A treatment.

Next, the impact of the *TP53* gene levels onto the wt HCT116 cell proliferation upon treatment with Epiblastin A was investigated (see **Figure 46**). The GI_{50} value was found to increase from the wt (*TP53*$^{+/+}$, GI_{50} = 11.81 ± 4.15 µM) to the mutant variants that either lack one allele (HCT116 *TP53*$^{+/-}$) or both *TP53* alleles (HCT116 *TP53*$^{-/-}$), whereby no GI_{50} value could be determined under the same experimental conditions.

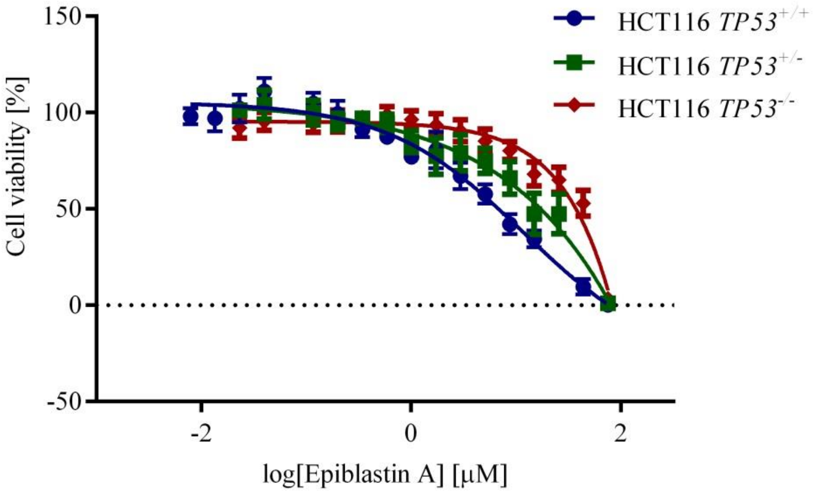


Figure 46. WST-1 cell proliferation assay of *TP53* mutant HCT116 cell lines upon Epiblastin A treatment for 24 h. Experimental details: 4000 cells of *TP53*$^{+/+}$ HCT116, *TP53*$^{+/-}$ HCT116 and *TP53*$^{-/-}$ HCT116 were initially plated overnight followed next day by addition of Epiblastin A at various concentrations. Afterwards, WST-1 reagent is added and the absorption at 450 nm is recorded. Cells treated with DMSO were considered 100 % viable. Data represent the average ± SD of three independent experiments performed in quadruplicate. Determined IC_{50} values ± SD were generated by fitting the data with four parameter nonlinear regression analysis using GraphPad Prism version 6.00 for Windows.

This behavior clearly indicates that *TP53*$^{-/-}$ HCT116 cells are more resistant to the growth inhibition mechanism enacted by Epiblastin A and suggests that the mode of action exerted by Epiblastin A is at least in part mediated via the TP53 axis. This is not surprising considering the literature reports linking CSNK1 kinase family to TP53 and other TP53 pathway components.[199]

In conclusion, manipulating the initial number of wt HCT116 cells plated indicated that increasing the degree of cell-cell contact in culture synergizes with the growth inhibition mechanism promoted by Epiblastin A. Moreover, the lack of *TP53* gene blocked this mechanism, suggesting that this gene is an integral part of Epiblastin's A mode of action.

4.11.3. Epiblastin A affects the growth of various cancerous and non-cancerous cell lines

To investigate the specificity of the growth inhibition mechanism upon Epiblastin A treatment, cytotoxicity studies were expanded on other cancerous (see **Figure 47**) and non-cancerous cell lines using WST-1 assay (see **Figure 48**).

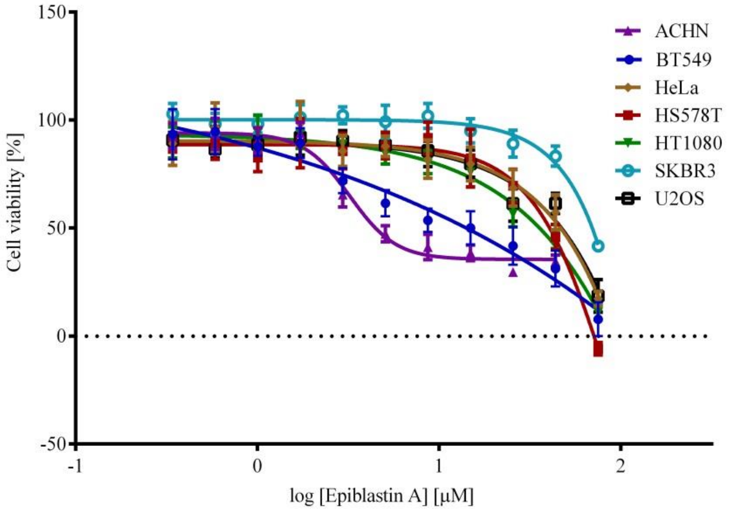


Figure 47. WST-1 cell proliferation assay of ACHN, BT549, HeLa, HS588T, HT1080, SKBR3 and U2OS cells upon Epiblastin A treatment for 24 h. Experimental details: cells were plated overnight followed next day by addition of Epiblastin A at various concentrations. After overnight incubation, WST-1 reagent is added and the absorption at 450 nm is recorded. Cells treated with DMSO were considered 100 % viable. Data represent the average ± SD of three independent experiments performed in quadruplicate. Determined IC_{50} values ± SD were generated by fitting the data with four parameter nonlinear regression analysis using GraphPad Prism version 6.00 for Windows.

As shown in **Figure 47**, Epiblastin A does not seemingly affect the proliferation of cancer cell lines such as HeLa, HS578T, HT1080, SKBR3 and U2OS. However, some degree of cytotoxicity was recorded in the case of BT549 ductal carcinoma cell line where an almost linear reduction of the mitochondrial dehydrogenases was observed upon increasing concentrations of Epiblastin A. An interesting behavior was observed in the case of ACHN renal adenocarcinoma cell line, where the presence of a resistant population was observed that was not affected even at high Epiblastin A concentration such as 75 μM.

However, the cell proliferation of non-cancerous cell lines (see **Figure 48**) such as human embryonic kidney 293 (HEK293) and the highly similar HEK293T were affected to a greater extent than the cancerous cell lines (GI_{50} = 37.63 ± 5.48 μM for HEK293 and GI_{50} = 8.54 ± 0.59 μM for HEK293T). This observation suggests an unspecific growth inhibition mechanism exerted by Epiblastin A. In addition, the proliferation of MCF10A, an epithelial cell line derived from breast, was moderately inhibited (GI_{50} = 24.55 ± 0.94 μM).

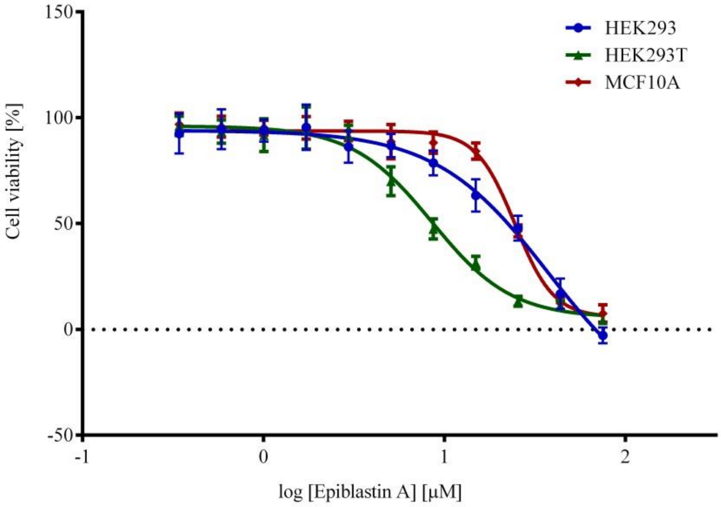


Figure 48. WST-1 cell proliferation assay of HEK293, HEK293T and MCF10A cells upon Epiblastin A treatment for 24 h. Experimental details: cells were plated overnight followed next day by addition of Epiblastin A at various concentrations. After overnight incubation, WST-1 reagent is added and the absorption at 450 nm is recorded. Cells treated with DMSO were considered 100 % viable. Data represent the average ± SD of three independent experiments performed in quadruplicate. Determined IC_{50} values ± SD were generated by fitting the data with four parameter nonlinear regression analysis using GraphPad Prism version 6.00 for Windows.

In conclusion, Epiblastin A's mode of action is not specific to cancer cell lines. Inhibition of the cell proliferation was significantly detected for non-cancerous cell lines compared to cancerous cell lines, suggesting an unspecific growth inhibition effect.

5. Discussion

Stem cell research can have a major impact on regenerative medicine in areas such as cell replacement therapies,[102,103] disease modeling[106] and tissue engineering.[104,105] One important prerequisite to exploiting the full clinical potential of stem cells is accessibility of a large amount of high quality embryonic stem cells. These cells are normally isolated from the inner cell mass of a developed embryo and maintained *in vitro* under defined culture conditions. Because of the ethical issues related to embryo manipulation for experimental purposes, researchers have developed novel protocols to access embryonic stem (like) cells. Following on the initial contributions from John B. Gurdon showing that cellular fate can be reset by means of nuclear reprogramming,[35] Yamanaka et al.[39] succeeded to generate induced pluripotent stem cells (iPSCs) from fully differentiated cells by overexpressing a cocktail of four transcription factors, *Oct4*, *Sox2*, *Klf4* and *Myc*. These cells are hardly distinguishable from the original ESCs in terms of morphology, culture conditions, expression of pluripotency-associated markers, differentiation potential, chimera contribution and other parameters.[39,40] However, deep molecular investigations revealed several drawbacks of iPSCs such as immunogenicity issues, re-expression of transgenes and tumorigenic potential due to the presence of *Myc* in the reprogramming cocktail. These disadvantages hamper the use of iPSC technology in clinical context. To overcome these shortcomings, non-integrative methods need to be developed that completely avoid the introduction of transgenes into the host’s genome. However non-integrative methods developed to date, are not optimal for clinical use due to relatively low yields of producing pluripotent stem cells.

One of the most appealing non-integrative methods relies on the use of small molecules. Such compounds were successfully employed in various stem cell research areas such as reprogramming of the cellular fate, differentiation towards various lineages and self-renewal of pluripotent stem cells. For example, small organic molecules are now an integrative part of various self-renewal protocols that robustly and reproducibly allow passaging of (pluripotent) stem cells *in vitro*. Additionally, many small molecules boost the generation of pluripotent stem cells in various contexts by either synergizing with the overexpressed transcription factors or even acting as surrogates for these transgenes. This allowed the continuous reduction of the number of transgene insertion in fully differentiated cells to reach an embryonic stem-like state. Recently, chemical reprogramming of fully differentiated mouse cells to mESCs was achieved by small molecules only, albeit with low yields.[94] This emphasized that the proper manipulation of the dosage and order of addition can successfully generate embryonic stem-like cells entirely avoiding any genetic intervention into the host's genome.

Due to the major role played by kinases in various aspects of stem cell biology, it is not surprising that kinase inhibitors proved to be active ingredients in various reprogramming contexts. However, most of the kinase inhibitors used to date were developed to fulfill the goals of other research areas such as oncology and repurposed for use is stem cell research. Thus, there is an obvious lack of small molecules that are particularly optimized in a stem cell based context. Moreover, there is an unmet need of novel chemical entities that would better address the major bottlenecks before clinical onset, in particular stem cell heterogeneity and low production yields.

As a model for the study of reprogramming, stem cells derived from the post-implantation (epiblast) developmental stage were chosen, i.e. epiblast stem cells (EpiSCs).[15,16] These cells are characterized by primed pluripotency[119] and are molecularly and functionally different than naïve ESCs. One major difference between mEpiSCs and mESCs is that the latter do contribute to chimera formation compared to mEpiSCs. Literature reports have shown that the mEpiSC population is heterogeneous[132] as assessed by the expression of *Oct4*, a crucial pluripotency associated marker. As a result, two distinct subpopulations were identified. On one hand, an Oct4-GFP-positive cell population was identified that represents less than 1 % from the total amount of cells and was termed early-stage mEpiSCs. On the other hand, 99 % of the remaining cells are Oct4-GFP-negative, do not express *Oct4* at detectable levels and were designated late-stage mEpiSCs. The two populations are distinct as indicated by the chimera formation capacity. The GFP-positive population contributes to chimera although with very low yields, whereas the GFP-negative population does not, suggesting that the two populations represent slightly different developmental stages.

Primed mEpiSCs can be converted to naïve mESCs, either by genetic intervention, chemical means or even a simple switch in the culture conditions.[136,137,138,139,140,141,142] However, in all cases an unsorted, i.e. heterogeneous mEpiSC population, was used. Thus, it might be possible that only the Oct4-GFP-positive cells were converted under the respective experimental conditions. To completely rule out this possibility, it has been previously shown that the commercialized drug Triamterene (**TR**) bearing a pteridine scaffold is able to reprogram late-stage mEpiSCs to mES-like cells.[148,149] In-depth characterization revealed that

TR first converts the initial population to a pre-ESC-like state that exhibits most of the naïve pluripotency hallmarks but lacks the ability to contribute to chimera formation. Chimera competent cells could only be generated if the pre-ES population was cultured in the presence of a known MEK1 inhibitor, PD0325901 (**PD**).

The current work reports on the optimization of the previously identified tool compound Triamterene (**TR)** through synthesis of a compound collection of closely related analogs. Their biological activity was evaluated in combination with **PD**, in a stem cell-based reprogramming assay. The principle of the assay relies on the reactivation of *Oct4* expression, the most important pluripotency-associated marker. This event was monitored through a fluorescent *Oct4*-GFP construct and quantified by means of flow cytometry.

The biological investigation of the **TR**-related analogues provided a detailed picture of the structural requirements needed for the reprogramming of late-stage mEpiSCs. Even small alterations of the parent **TR** derivative have a strong impact on the phenotype. First, for an optimal reprogramming activity, the phenyl ring at position C6 featuring a defined substitution pattern needs to be maintained. Second, the derivatization of the meta- and ortho-positions with specific groups is allowed whereas any manipulation of the para-position led to reduced or even abolished biological activity. Moreover, the amino-groups in positions C2 and C7 proved to be essential as well, since their replacement with various substituents was detrimental to the reprogramming activity.

Analysis of the pteridine library in the reprogramming assay yielded five derivatives whose biological activity was improved compared to the initial hit **TR**. These compounds were collectively designated Epiblastins. The most potent compound was termed Epiblastin A whereas the remaining four derivatives are named according to their descreasing activity with respect to the initial compound, i.e. Epiblastin B-E. Since the biological activity and mode of action were already throughly characterized for Epiblastin A,[149] the current work focused on the in-depth characterization of Epiblastin B-E. For these derivatives, additional experiments were employed to prove the reprogramming activity. First, the Epiblastin B-E reprogramming activities were analyzed together with Epiblastin A and included an authentic mESC cell line for which the endogenous Oct4-GFP expression was quantified. These derivatives reproducibly converted late-stage mEpiSCs to Oct4-GFP-positive cells with yields ranging from 38 % for Epiblastin E to approximately 70 % for Epiblastin A. Therefore, between 38 % and 70 % of the initial starting population reactivated the expression of the Oct4-GFP reporter and were converted to a naïve state. In comparison, the mESC cell line was evaluated to 93 %, whereas **TR** converted only 31 % from the initial population. Moreover, by means of immunofluorescence experiments, the expression of *Zpf42* (also known as *Rex1*) could be detected in the readily reprogrammed Oct4-GFP-positive cells. This marker is associated with the naïve state exclusively and proves that the cellular state obtained by treatment of Epiblastin A-E and **PD** is indeed embryonic stem-like. Taking together, Epiblastins A-E represent important tool compounds for chemical reprogramming of partial or fully differentiated cells towards naïve pluripotency.

Although **TR** was previously reported to target the epithelial sodium channel (ENaC)[150] and dihydrofolate reductase (DHFR),[241,242] this compound has been barely investigated for inhibition of kinase targets. A single study highlights that **TR** fulfills the ”2-0” kinase-likeness rule[243] which allows to distinguish small molecules or fragments that inhibit protein kinases from compounds that do not interact with this class of enzymes. If the sum of heteroaromatic nitrogen atoms and heteroaromatic NH groups embedded within the compound structure is greater than 2, than the compound is likely to inhibit protein kinases. The study found that SGK1, p70 S6K and PAK1 kinases were completely inhibited by **TR** at a concentration of 33 μM, but no dose-dependent studies have been reported so far. In addition to the literature reports, **TR** and Epiblastin A have been shown to specifically inhibit the PI3K and CSNK1 kinase families in a profiling study of more than 100 kinases.[149] The contribution of PI3K to the reprogramming mechanism was devalidated using two structurally unrelated inhibitors, LY294002 and AS605240. However, the pan-CSNK1 inhibitor D4476 was able to generate Oct4-GFP colonies suggesting that blocking the CSNK1 activity mediates the cellular transition mechanism. Among the CSNK1 family, the genetic ablation of CSNK1A1 and CSNK1E could reproduce the phenotype. In particular, knock-down of the CSNK1A1 isoenzyme by means of shRNA showed the highest percentage of Oct4-GFP expressing cells among the CSNK1 isoforms tested.

Since CSNK1 is involved in the reprogramming of late-stage mEpiSCs, a target-based approach was applied to characterize the inhibitory activity of the synthesized compound library for the CSNK1A1/D/E isoforms. The collection was initially profiled for inhibition of CSNK1D/E isoenzymes using an in-house developed assay based on the KinaseGlo® reagent. This robust and reproducible assay showed very good Z’ factor values. Using this platform, a ATP competitive inhibition by Epiblastin A was determined for both CSNK1D/E isoenzymes as concluded from Lineweaver-Burk plots and IC_{50} determinations at different ATP concentrations. Moreover, selected examples were analyzed at Reaction Biology (Pasadena, USA) and analyzed for inhibition of CSNK1A1.

Profiling of the pteridine collection allowed the delineation of clear structure-activity relationship parameters to achieve both potency and / or selectivity across the CSNK1A1/D/E isoforms. Protein tree analysis indicates the high structural similarity of CSNK1D and CSNK1E, substantiating the difficulty of achieving selectivity across these two isoforms. On the contrary, the modulation of the substitution pattern on the phenyl ring at position C6 allows for the generation of isoform-selective inhibitors. For example, the chlorine and bromine derivatization in the meta-position as shown for Epiblastin A and derivative **15** renders CSNK1D and CSNK1A1/D inhibitors, respectively. The replacement of the phenyl ring with naphthalene unit as shown for Epiblastin C switches the selectivity towards CSNK1E. As a result, Epiblastin A and Epiblastin C respresent selective CSNK1D/CSNK1E inhibitors, outmatching commercially available inhibitors such as PF-670462 and PF-4800567.[214] Thus, Epiblastin A and Epiblastin C are novel tools to investigate CSNK1-mediated processes in an isoform-selective fashion. For example, inhibition of this kinase family led to therapeutic gain in diseases such as acute myeloid leukemia (AML),[244,245] amyotrophic lateral sclerosis (ALS)[207] and various types of cancers.[246,247,248,249]

The structure-activity relationships for the CSNK1A1/D/E isoenzymes corroborate well with results of the stem cell reprogramming assay. In both cases, the derivatization of the para-position of the C6 phenyl ring was detrimental to activity, whereas specific meta- or ortho-substitutions led to higher activity. There are few exceptions worth mentioning. The derivative containing a bromine atom in the meta-position of the phenyl ring at position C6 inhibits CSNK1A1/D enzymatic activity. Although the IC_{50} values are lower than for **TR**, the reprogramming activity is almost identical to **TR**. Conversely, the derivative containing two chlorine atoms in ortho- and ortho'-positions was inactive for inhibition of CSNK1A1/D/E isoenzymes but two times more potent than **TR** in the cell-based assay. These observations might point out the following assumptions: (1) there are additional targets involved other than CSNK1A1/D/E and their modulation might be crucial for generating the phenotype; (2) reprogramming can occur independently of CSNK1A1/D/E via another mechanism than previously disclosed for Epiblastin A. The existence of multiple pathways for reaching a certain cellular state is not surprising when considering the cellular plasticity[250,251] and heterogeneity that characterizes stem cells.[124] Moreover, several studies reported that the modulation of various targets by genetic and / or chemical means at the EpiSC stage can convert these cells to naïve mES-like cells (see section *1.1.10. Reprogramming mEpiSCs to mES-like cells – state of the art*).

The target engagement of CSNK1A1/D/E isoforms by Epiblastin A was demonstrated in ATP-depleted cell lysate derived from wt HCT116 (wt refers to the parental cell line containing both alleles of *TP53*, i.e. *TP53*$^{+/+}$) colorectal carcinoma cell line. This cell line was chosen because CSNK1A1/D/E could be detected at the protein level and treatment with Epiblastin A for 48 h significantly affected the cell proliferation as evaluated by the WST-1 assay. To address the interaction of Epiblastin A with the CSNK1 kinase family the commercially available ActivX^{TM} ATP probe was employed. This has successfully been applied before to ATP-binding proteins, collectively known as the purinome. This probe features an ATP molecule and contains a labile ester group at the γ-phosphate position that terminates in a desthiobiotin tag connected via a short linker. This probe effectively engages the nucleotide binding pocket of various classes of ATP binding proteins most of which feature one or even two lysine residues in close proximity of the nucleotide binding pocket. Thus, the lysine side chain reacts with the mixed anhydride, the engaged protein is labeled and this can be enriched using streptavidin beads. When profiling an ATP competitive inhibitor, the derivative can compete with the ActivX^{TM} ATP probe and thus the degree of labeling is diminished as compared to the treated sample. The labeling extent can be evaluated by either immunoblotting using a specific antibody when the respective protein target is known or by means of mass spectrometry. Engagement of CSNK1A1/D/E isoforms by Epiblastin A was initially evaluated by means of immunoblotting. This experiment revealed that concentrations higher than 10 μM were necessary to observe a significant protection effect against labeling. The only exception was CSNK1D, where 10 μM produced a noticeable effect, in line with the IC_{50} values reported *in vitro*. Thus, Epiblastin A engages CSNK1A1/D/E isoenzymes in cellular lysates and shows selectivity for CSNK1D.

To reveal ATP-binding protein targets that might additionally be inhibited by Epiblastin A in HCT116 cell lysates, a chemical proteomics approach was employed. After incubation with Epiblastin A and the ActivX^{TM} ATP probe, the labeled proteins are enriched using streptavidin coated beads, washed, the bound proteins were digested and the resulting peptides analyzed by means of LC-MS. This procedure was applied in different formats by varying the incubation time of Epiblastin A and the ActivX^{TM} ATP probe from 10 to 60 min. An average of 1440 proteins per experiment was identified, out of which 25 % have been reported to bind ATP by the molecular function gene ontology (GOMF) analysis.

Approximately 40 % from the ATP binding proteins enriched were kinases covering approximately 27 % of the entire kinome. As a result, the statistically relevant presence of CSNK1A1 was confirmed in the protein list generated from all four chemical proteomics experiments. Gratifyingly, the same isoform was reported in a previous study to be the main target driving the reprogramming of late-stage mEpiSCs. No other CSNK1 isoenzyme could be identified in the protein list during the experimental procedures. This does not exclude the interaction with Epiblastin A, but rather points out a lower abundance of the other isoenzymes in HCT116 cell lysate.

In addition to CSNK1A1, STK10 (Serine / threonine-protein kinase 10) was identified in all experiments as main target possibly engaged by Epiblastin A. By mean of IC_{50} determination *in vitro* and target engagement in HCT116 cell lysates using immunoblotting, STK10 was confirmed as an interaction partner of Epiblastin A. STK10 is the human homologue of murine LOK (Lymphocyte oriented kinase)[252] and is a serine / threonine kinase highly expressed in lymphocytes where it fulfills various biological roles.[253,254] Recently it has been reported that STK10 functions as a tumor suppressor, because missense mutations within the STK10 sequence led to anti-apoptotic activity in peripheral T-cell lymphoma (PTCL) cellular model.[255] Moreover, STK10 acts as a polo-like kinase kinase regulating the activation status of polo-like kinase 1 (PLK1).[239] In addition, STK10 selectively modulates the proliferation of Ewing's sarcoma cells.[256] Knockdown by means of siRNA (small interfering RNA) selectively induced apoptosis in model cell lines such as TC-32, TC-71, SK-ES-1 and RD-ES compared to normal fibroblasts.

The chemical proteomics experiments additionally identified with high statistical confidence two members of the phosphatidylinositol-5-phosphate (PtdIns5*P*)-4-kinases (PIP4Ks) kinase family, PIP4K2A (Phosphatidylinositol 5-phosphate 4-kinase type-2 alpha) and PIP4K2C (Phosphatidylinositol 5-phosphate 4-kinase type-2 gamma). This class of lipid kinases readily phosphorylates PtdIns5*P* to PtdIns(4,5)P_2 and mediates oxidative stress signaling. Therefore, the members of PIP4K family have been recently proposed as a novel class of anti-cancer therapeutic agents.[257,258] This family comprises three members, namely PIP4K2A, PIP4K2B and PIP4K2C and the deregulation of the enzymatic activity directly affects the accumulation of reactive oxygen species in mammalian cells. PIP4K2A is by far the most active isoenzyme from this series being 100-500 fold more active than the PIP4K2B and more than 1000 fold more potent than PIP4K2C.[259,260] PIP4K2A has been recently proposed as a novel target for treatment of amyotrophic lateral sclerosis.[261]

In addition to the primary list of targets presented above, other ATP-binding proteins could be identified. These appeared only once in each experimental condition employed and might represent only transient associations with Epiblastin A. For example, SLK (STE20-like serine / threonine-protein kinase) is highly related to STK10 containing an identical sequence within the ATP binding site. Literature reports suggest that STK10 together with SLK[239,262] modulate the enzymatic activity of polo-like kinase 1 (PLK1) in tissue culture. In addition, by sequence homology searches, PRKDC (known as DNA-PK)[263] has been identified as belonging to the phosphatidylinositol 3-kinase family (PI3K).[264,265] Studies have shown that the PI3K catalytic domain of PRKDC kinase phosphorylates a broad range of substrates but does not exhibit lipid kinase activity.[265] Similar to PI3K family, PRKDC is inhibited by high concentrations of Wortmannin.[266] Together with the previous results related to the possible inhibition of PIP4K by Epiblastin A and the *in vitro* studies against PI3K, it can be surmised that the pteridine scaffold inherently targets lipid kinase catalytic domains. Further studies might offer the possibility to develop specific inhibitors for lipid kinases.

With the exception of CSNK1A1 that was associated with the DNA-damage response in mouse ESCs[267] only two putative targets identified by the chemical proteomics experiment were previously reported in the stem cell literature. MAPK8 (known as JNK1) was previously reported as barrier in reprogramming.[268] Thus, it has been shown that JNK1 and the closely related JNK2 isoform directly phosphorylate Thr224 and Thr225 residues of *Klf4*, thus inhibiting its transcription and suppressing the transactivation activity. As a result, knockout JNK1$^{-/-}$ or JNK2$^{-/-}$ mouse embryonic fibroblasts (MEFs) yielded a significantly higher number of pluripotent stem cell colonies when reprogrammed with Yamanaka factors compared to the wt MEFs.

NRBP1 (Nuclear receptor binding protein 1)[269] could be a potential candidate for further investigations. The loss of this pseudokinase[270,271] was reported to activate the Wnt pathway and led to SALL4 accumulation.[272] SALL4 is a crucial transcription factor that either directly or indirectly modulates the gene network responsible for pluripotency establishment.[11,273,274] NRBP1 may be a tumor suppressor[269,275,276] and its high expression in associated with poor prognosis and marked tumor growth.

In conclusion, the chemical proteomics procedure based on the ActivX^{TM} ATP probe revealed potential novel targets addressed by Epiblastin A in cell lysates derived from the HCT116 colorectal carcinoma cell line. These proteins might synergize with the CSNK1 kinase inhibition and might explain the increased reprogramming efficiency of Epiblastin A compared to **TR**.

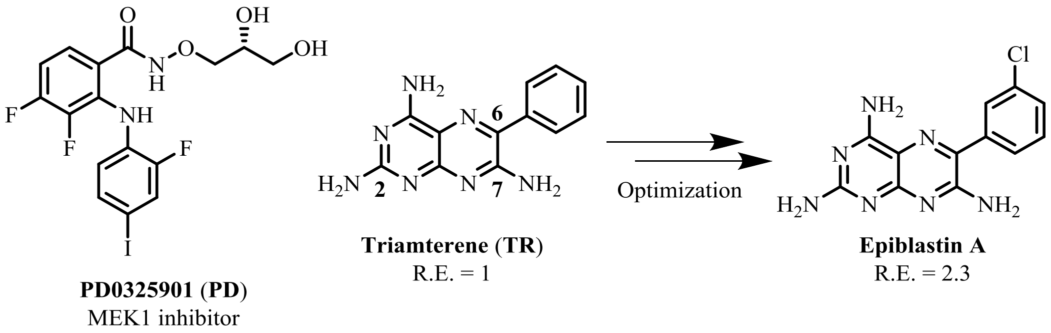
After determining additional binding partners of Epiblastin A in HCT116 cell lysate, this molecule was submitted to a profiling study *in vitro* comprising a panel of 123 kinases in the attempt to reveal new targets that might have been missed during the proteomics procedure likely due to low abundance. Single-point measurement followed by dose dependent analyses revealed a set of five additional kinases, namely BRSK1, EEF2K, EGFR, MKNK2 and RIPK2 with IC_{50} values ranging from 8.3 μM for EGFR to 45 μM for MKNK2. With the exception of EGFR, the other kinases exhibited IC_{50} values *in vitro* much higher compared to CSNK1A1. Interestingly, EGFR was recently found to play the role of a barrier in the transition from an intermediate pre-ESC state derived from mouse to an ES-like state. Tran et al.[277] found that removing the components of the EGF pathway facilitated this late-stage reprogramming by upregulating *Esrrb*, a known transcription factor that is a part of the pluripotency network in mouse. Interestingly, we have also previously shown that Epiblastin A treatment of late-stage mEpiSCs steadily activates the expression of *Esrrb*. Future experiments are required to address the possible involvement of EGFR inhibition in the context of reprogramming late-stage mEpiSCs by Epiblastin A.

In addition to the reprogramming capacity of Epiblastin A, this compound was investigated in the context of cancer cell research as well. Following this direction, it was proven that Epiblastin A blocks the proliferation of wt HCT116 colorectal carcinoma cell line as evaluated by the WST-1 assay. This assay evaluates the number of viable cells in culture by quantifying the enzymatic activity of mitochondrial dehydrogenases. The WST-1 assay suggested that prolonged incubation time is required for a significant inhibition of wt HCT116 cell growth. Similar conclusions were obtained by counting the number of viable cells directly in culture by means of Tryphan blue and CellTiter-Glo® cell viability assay. Tryphan blue specifically stains the nucleus of dead cells whereas CellTiter-Glo® quantifies the number of viable cells by measuring the cellular ATP content. Finally, a CellTox™ Green assay was employed to assess the membrane integrity of the cells upon Epiblastin A treatment directly in culture. The result of this analysis showed that upon 24 h treatment, only high concentrations of Epiblastin A, i.e. 50 μM produced significant cell membrane disruption. As expected, prolonged time significantly reduced the number of viable cells. Overall, these results point out that Epiblastin A significantly inhibits the cell proliferation of wt HCT116 cells only upon prolonged incubation time, namely 48 h. Moreover, testing the pan-CSNK1 inhibitor D4476 in the WST-1 assay for 24 h incubation, suggested that the growth inhibition mechanism exerted by Epiblastin A might not be due to CSNK1 inhibition. The growth inhibition was shown to be affected by parameters such as initial number of cells plated and the *TP53* gene status. Increasing the initial number of wt HCT116 cells plated from 4,000 to 12,000 considerably decreased the GI_{50} value as evaluated by the WST-1 assay. Increasing the cell number even further to 20,000 cells did not significantly alter this value suggesting a saturation effect. This might indicate that the extent of cell-cell contact might mediate the cytotoxic mechanism exerted by Epiblastin A. The lack of *TP53* gene seemed to have a negative effect on the cell growth inhibition mechanism exerted by this derivative. Whereas the presence of both *TP53* alleles in the HCT116 $TP53^{+/+}$ cell lines yielded a GI_{50} value of 12 µM, no GI_{50} value could be determined in the case of the mutant cell lines lacking one (HCT116 $TP53^{+/-}$) or both *TP53* alleles (HCT116 $TP53^{-/-}$). This result suggests the inhibition of the HCT116 cell proliferation mediated by Epiblastin A requires *TP53* gene product. This is not surprising when considering that the CSNK1 kinase family members directly phosphorylate amino acid residues within the TP53 sequence and within its partners.[199,278,279,280]

After testing the wt HCT116 cell line, the cytotoxic properties of Epiblastin A were investigated on other cancerous and non-cancerous cell lines. No significant growth inhibition effect was observed for HeLa, HS578T, HT1080, SKBR3 and U2OS cell lines. Some degree of cytotoxicity was observed in the case of BT549 ductal carcinoma cell line where an almost linear reduction of the WST-1 signal was observed upon increasing concentrations of Epiblastin A. An interesting behavior was obsverved in the case of ACHN renal adenocarcinoma cell line, where the presence of a resistant population was observed that was not affected even at 75 μM Epiblastin A. Surprisingly, non-cancerous cell lines such as HEK293, HEK293T and MCF10A were affected to a much greater extent than the corresponding cancerous counterparts. These results suggest that the cell growth inhibition mechanism mediated by Epiblastin A is rather unspecific across different cell lines.

6. Summary

In the course of this work, an in-depth chemical biology investigation has been performed to study the reprogramming of partially differentiated cells termed late-stage mouse epiblast stem cells (mEpiSCs) to embryonic stem (ES)-like cells upon treatment with small molecules. Triamterene (**TR**) has been previously reported to convert the starting cell population to intermediate embryonic stem (ES)-like cells (designated pre-ESCs) alone or to ES-like cells when treated in combination with the MEK1 inhibitor PD0325901 (**PD**) (see chemical structures in **Figure 49**).



PD0325901 (**PD**)
MEK1 inhibitor

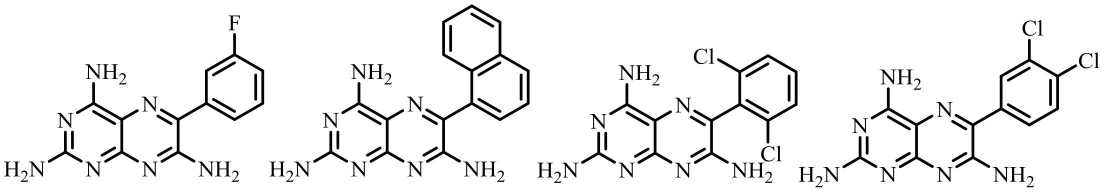
Triamterene (**TR**)
R.E. = 1

Epiblastin A
R.E. = 2.3

Figure 49. Chemical structures of PD0325901 (**PD**), Triamterene (**TR**) and Epiblastin A. R.E. is defined as the reprogramming efficiency; for the initial hit **TR**, R.E. was set to 1.

To relate chemical structure with biological activity, a compound collection based on the **TR** scaffold was synthesized bearing various substituents at positions C2, C6 and C7 (see above) following procedures already established in the literature. For example, mono- and disubstituted derivatives on the phenyl ring at position C6 can be synthesized in a single step by reacting 5-nitroso-2,4,6-triaminopyrimidine with commercially available phenylacetonitriles. Analogously, pteridines functionalized at position C2 can be generated from various 2-substituted-5-nitroso-4,6-diaminopyrimidines.

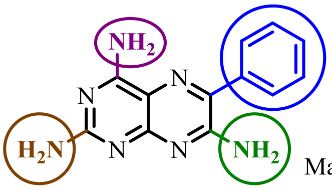
The biological investigation of the compound collection was performed by quantifying the selective re-expression of the Oct4-GFP construct, a crucial pluripotency-associated marker, in the reprogrammed cells. The structure activity relationship study defined clear structural features required for increased reprogramming activity with respect to the initial derivative **TR**. Thus, it could be shown that small changes in the structure of the parent compound have dramatic effects on the extent of reprogramming late-stage mEpiSCs. Epiblastin A, a meta-chloro phenyl-substituted **TR** analog (see chemical structure in **Figure 49**) was identified as the most active compound that reproducibly achieved reprogramming efficiency of up to 2.3 fold greater than the parental compound **TR**. In addition to Epiblastin A, four other pteridine derivatives termed Epiblastin B-E were discovered that exhibited increased conversion rates compared to **TR** (see chemical structures of Epiblastin B-E in **Figure 50**).



Epiblastin B R.E. = 1.9 **Epiblastin C** R.E. = 1.8 **Epiblastin D** R.E. = 1.8 **Epiblastin E** R.E. = 1.2

Figure 50. Chemical structures of Epiblastin B-E. R.E. is defined as the reprogramming efficiency; for the initial hit **TR** was set to 1.

Because casein kinase 1 (CSNK1) family was previously identified to be mediate the cellular transition mechanism, pteridine derivatives were investigated for inhibiting the enzymatic activity of CSNK1 alpha (A1), CSNK1 delta (D) and epsilon (E). Clear structure-activity relationships could be delineated and revealed structural modifications that interfere with the CSNK1A1/D/E kinase activity (see **Figure 51**). For example, simple substitutions at position C6 as shown for Epiblastin A and Epiblastin C yielded selective CSNK1D/CSNK1E inhibitors, respectively.



Mandatory for CSNK1D/E activity

Specific substitutions render either CSNK1D or CSNK1E inhibitors

Mandatory for CSNK1D/E activity. Replacement with -H renders selectivity towards CSNK1E

Mandatory for CSNK1D/E activity

Figure 51. Conclusions of the structure-activity relationship study for pteridine derivatives derived from enzymatic studies CSNK1D/E activity.

The Lineweaver-Burk analyses suggested that Epiblastin A is an ATP-competitive inhibitor of CSNK1D/E isoenzymes. This conclusion was supported by the shift of the half-maximal inhibitory concentration (IC_{50}) to higher values by increasing the ATP concentration in the enzymatic reaction catalyzed by CSNK1D/E.

The crystal structure of CSNK1D in complex with Epiblastin A or derivative **15** which contain a chlorine and bromine atom in the meta position of the C6 phenyl ring, respectively, furnished important insights on the mode of binding for this class of compounds. These derivatives dock in the nucleotide binding pocket by means of four hydrogen bonds with the amino acid backbone of the hinge region. Moreover, the chorine / bromine atom in the meta position orient towards the hydrophobic cavity created by the amino acid residues of Met80 and Met82. The crystal structure of CSNK1D in complex with the two inhibitors provides insights into how to design novel pteridine derivatives with improved potency and / or selectivity against CSNK1D isoform.

Target engagement of CSNK1A1/D/E isoenzymes by Epiblastin A in cell lysates was investigated using the ActivX^{TM} ATP probe, a reagent that specifically interacts with ATP-binding proteins, collectively designated as the purinome. This agent is suited for activity-based profiling studies, where it competes with the ATP competitive inhibitor under investigation for the nucleotide binding pocket of the addressed target in complex protein mixture. The degree of interaction can easily be assessed by means of immunoblotting if the target of interest is known beforehand. Using this strategy significant engagement of Epiblastin A with CSNK1A1/D/E isoenzymes was detected in HCT116 cell lysate confirming the interaction of this derivative with the CSNK1 family in complex protein mixture.

Moreover, by means of a chemical proteomics strategy using the same reagent, CSNK1A1 was confirmed the main target of Epiblastin A in HCT116 cell lysate. In addition, serine / threonine-protein kinase 10 (STK10), phosphatidylinositol 5-phosphate 4-kinase type-2 alpha (PIP4K2A) and phosphatidylinositol 5-phosphate 4-kinase type-2 gamma (PIP4K2C) kinases were identified as interaction partners of Epiblastin A. Among these targets, solely the interaction of Epiblastin A with STK10 was in-depth investigated in the current study and subsequently proven *in vitro* with purified enzyme and in cell lysate using the ActivX^{TM} ATP probe.

To reveal other kinases potentially addressed by Epiblastin A, an *in vitro* kinase profiling study comprising 123 kinases was performed which additionally revealed brain-selective kinase 1 (BRSK1), eukaryotic elongation factor 2 kinase (EEF2K), epidermal growth factor receptor (EGFR), MAP kinase-interacting serine / threonine-protein kinase 2 (MKNK2) and receptor-interacting serine / threonine-protein kinase 2 (RIPK2) as potential targets of Epiblastin A. However, the half-maximal inhibitory concentration values for inhibiting these targets were significantly higher than for CSNK1A1 isoenzyme with the exception of EGFR, suggesting that only the EGFR inhibition might be relevant in the context of reprogramming late-stage mEpiSCs. These studies potentially revealed other targets engaged by Epiblastin A in HCT116 cell lysate that might represent potential reprogramming barriers during the cellular transition process.

Thus, chemical biology approaches can have a tremendous impact in stem cell research areas such as cell dedifferentiation mediated by small molecules, i.e. chemical reprogramming. The discovery of Epiblastin A-E provides the stem cell community new tools that avoid the genetic intervention and reprogram stem cells derived from the later stages of the embryo development, i.e. postimplantation stage, to ES-like cells. Moreover, the current study identified interaction partners of Epiblastin A whose inhibition might contribute to the generation of high quality pluripotent stem cells with high yields, thus bringing stem cell-based technologies closer to clinical applications.

7. Outlook

The results highlighted in the present study could be extended in various directions. For example, developing new methods to access substituted pteridine libraries would have a great impact on the various phenotypes associated with these compounds. Generating novel Epiblastin A-based derivatives embodying various substituents at positions C2, C4 and C7 could improve even further the reprogramming activity by iterative cycles of synthesis and biological evaluation. Not only potency can in principle be improved but also the cytotoxicity reported for these derivatives in the biological context.

Some of the compounds generated in this report already showed good selectivity profile across the CSNK1 kinase family. Another direction to follow would be the improvement of the potency and selectivity in an attempt to develop isoform-specific inhibitors. Along these lines, Epiblastin A and Epiblastin C are already good candidates to start medicinal chemistry campaigns oriented towards these goals. Moreover, the crystal structure of Epiblastin A already indicates derivatization points that can be further exploited. The improved inhibitors could then be tested in various disease models where CSNK1 kinase activity plays an essential role such as acute myeloid leukemia, amyotrophic lateral sclerosis and various types of cancers.

Another interesting aspect of these derivatives that can be exploited is the observation that pteridine compounds can reprogram late-stage mEpiSCs alone to an intermediate pre-ESC like state. This particular trapped population offers the starting platform to test various stem cell differentiation protocols. Thus, it might be possible that this reprogrammed state is biased to differentiate to cells belonging to specific lineages with higher propensity than ES-like cells.

Thus far we have shown that the combined action of Epiblastin A and **PD** successfully converts late-stage mEpiSCs to ES-like cells with high yields. This represents the first step towards the development of a rational strategy to achieve chemical reprogramming with high yields starting from a fully differentiated state. The next logical step would be to identify small molecules that are able to convert somatic cell populations into a state that resembles the late-stage epiblast. Moreover, it would be interesting to test if Epiblastin A-E mode of action could promote naïve pluripotency in hESCs, since it is nowadays well accepted that these cells resemble mEpiSCs in many cellular and functional aspects.

Another important issue to optimize is related to the chemical proteomics protocol. Although this so far has proven to be relatively robust in identifying ATP binding proteins from entire proteomes, it still requires individual target engagement confirmation experiments using immunoblotting for the proposed target addressed by the small molecule under investigation. Therefore an increased throughput is required to rapidly delineate between positive and negative protein target hits. This could easily be achieved by a simple change of the original protocol. For example, after incubation with the compound of interest and the ActivX^{TM} ATP probe, labeled proteins are directly digested instead of subsequent enrichment followed by digestion. After this step, the labeled peptides are collected using streptavidin coated beads, washed, boiled and identified by means of LC-MS. This particular change of the original protocol could foster the identification of the exact position of the lysine amino acid residue within the protein sequence that has been labeled by the ActivX^{TM} ATP probe. By mapping this lysine residue in the structure of the ATP binding proteins one could simply delineate between specific and unspecific labeling effects since only the lysine residue located in the vicinity of the binding pocket should in principle be targeted.

The identification of other interaction partners of Epiblastin A and the related analogues in lysates derived from mEpiSCs would offer a better understanding of the reprogramming process promoted by these derivatives. So far the experiments have been performed in cell lysates obtained from wt HCT116. It would be beneficial to check if the same targets are being addressed in lysates derived from mEpiSCs. The protein targets revealed by the chemical proteomics study performed in wt HCT116 cell line might represent novel barriers in the reprogramming of late-stage mEpiSCs. These proteins might synergize with CSNK1 kinase inhibition and explain, for example, the increased activity of Epiblastin A compared to the initial hit **TR**.

8. Appendix

Protein Names	10 min / 10 min	10 min / 60 min	60 min / 10 min	60 min / 60 min
STK10	7,95	2,13	7,66	3,10
PIP4K2A	1,82	2,07	2,10	1,76
CSNK1A1	3 / 0	3 / 0	2,20	1,59
SLK	1,67		1,84	1,44
PIP4K2C		1,52	1,40	1,50
DYNC1H1	20,41			
PRKDC	6,11			
MAP2K4			1,61	
ILK			1,54	
GNE	1,49			
GART		1,26	1,29	
PCCB				1,35
TAOK3			1,24	
ABCF1		1,22		
MVD		3 / 0		
PRKACB		3 / 0		
CHEK2	3 / 0			
NRBP1	3 / 0			
TNK1	3 / 0			
MAPK8		3 / 0		
TP53RK				3 / 0

Table 11. Total list of ATP-binding proteins identified by the chemical proteomics method employing the ActivX^{TM} ATP probe. The numbers represent the ratio between the peptide signal intensity detected in the DMSO and Epiblastin A (100 µM) treated samples, respectively after incubation with the ActivX^{TM} ATP probe. The higher the ratio, the higher the capacity of Epiblastin A to protect the ATP-binding protein against labeling with the ActivX^{TM} ATP reagent. 3 / 0 = suggest the case where the protein target could be identified in the DMSO control but not a single time in the Epiblastin A treated sample.

9. Abbreviations

ALS	Amyotrophic lateral sclerosis
AML	Acute myeloid leukemia
AP	Alkaline phosphatase
ATP	Adenosine triphosphate
BSA	Bovine serum albumin
CRISPR / Cas9	Clustered Regularly Interspaced Short Palindromic Repeats / Cas9
CSNK1	Casein kinase 1
DMSO	Dimethylsulfoxide
DNMT	DNA methyltransferase
dpc	Days post conception
EGC	Embryonic germ cells
ENaC	Epithelial sodium channel
EPL stem cells	Early primitive ectoderm-like stem cells
FAB stem cells	bFGF, Activin and BIO-derived stem cells
FACS	Fluorescence activated cell sorting
FDA	US Food and Drug Administration
FGF	Fibroblast growth factor
F.U.	Fluorescence units
GFP	Green fluorescent protein
GI_{50}	Half-maximal growth inhibitory concentration
GOMF	Molecular function gene ontology
HCSs	Hematopoietic stem cells
HDAC	Histone deacetylase
hESCs	Human embryonic stem cells
hiPSCs	Human induced pluripotent stem cells

IC_{50}	Half-maximal inhibitory concentration
ICM	Inner cell mass
iPSCs	Induced pluripotent stem cells
LC-MS	Liquid chromatography mass spectrometry
LFQ	Label free quantification
LIF	Leukemia inhibitory factor
LOPAC	Library of Pharmacologically Active Compounds
MEF	Mouse embryonic fibroblast
mEpiSCs	Mouse epiblast stem cells
mESCs	Mouse embryonic stem cells
MS	Mass spectrometry
NIH	National Institute of Health
PD	PD0325901
PGC	Primordial germ cells
PI3K	Phosphatidylinositol-3-kinase
R.L.U.	Relative light units
RNAi	RNA interference
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
shRNA	Small hairpin RNA
SSEA-1	Stage-specific embryonic antigen 1
TR	Triamterene
Xa	X-chromosome activation
XEN stem cells	Extra-embryonic endoderm stem cells
Xi	X-chromosome inactivation

10. References

1. Wobus, A. M.; Boheler, K. R. *Physiological Reviews* **2005**, *85*, 635.
2. Jaenisch, R.; Young, R. *Cell* **2008**, *132*, 567.
3. Boyer, L. A.; Lee, T. I.; Cole, M. F.; Johnstone, S. E.; Levine, S. S.; Zucker, J. P.; Guenther, M. G.; Kumar, R. M.; Murray, H. L.; Jenner, R. G.; Gifford, D. K.; Melton, D. A.; Jaenisch, R.; Young, R. A. *Cell* **2005**, *122*, 947.
4. Loh, Y.-H.; Wu, Q.; Chew, J.-L.; Vega, V. B.; Zhang, W.; Chen, X.; Bourque, G.; George, J.; Leong, B.; Liu, J.; Wong, K.-Y.; Sung, K. W.; Lee, C. W. H.; Zhao, X.-D.; Chiu, K.-P.; Lipovich, L.; Kuznetsov, V. A.; Robson, P.; Stanton, L. W.; Wei, C.-L.; Ruan, Y.; Lim, B.; Ng, H.-H. *Nature Genetics* **2006**, *38*, 431.
5. Thomson, M.; Liu, Siyuan J.; Zou, L.-N.; Smith, Z.; Meissner, A.; Ramanathan, S. *Cell* **2011**, *145*, 875.
6. Pera, M. F.; Tam, P. P. L. *Nature* **2010**, *465*, 713.
7. Niakan, K. K.; Davis, E. C.; Clipsham, R. C.; Jiang, M.; Dehart, D. B.; Sulik, K. K.; McCabe, E. R. B. *Molecular Genetics and Metabolism* **2006**, *88*, 261.
8. Ying, Q.-L.; Nichols, J.; Chambers, I.; Smith, A. *Cell* **2003**, *115*, 281.
9. Niwa, H.; Ogawa, K.; Shimosato, D.; Adachi, K. *Nature* **2009**, *460*, 118.
10. Heng, J.-C. D.; Feng, B.; Han, J.; Jiang, J.; Kraus, P.; Ng, J.-H.; Orlov, Y. L.; Huss, M.; Yang, L.; Lufkin, T.; Lim, B.; Ng, H.-H. *Cell Stem Cell* **2010**, *6*, 167.
11. Zhang, J.; Tam, W.-L.; Tong, G. Q.; Wu, Q.; Chan, H.-Y.; Soh, B.-S.; Lou, Y.; Yang, J.; Ma, Y.; Chai, L.; Ng, H.-H.; Lufkin, T.; Robson, P.; Lim, B. *Nature Cell Biology* **2006**, *8*, 1114.
12. Martin, G. R. *Proceedings of the National Academy of Sciences* **1981**, *78*, 7634.
13. Evans, M. J.; Kaufman, M. H. *Nature* **1981**, *292*, 154.
14. Thomson, J. A.; Itskovitz-Eldor, J.; Shapiro, S. S.; Waknitz, M. A.; Swiergiel, J. J.; Marshall, V. S.; Jones, J. M. *Science* **1998**, *282*, 1145.
15. Brons, I. G. M.; Smithers, L. E.; Trotter, M. W. B.; Rugg-Gunn, P.; Sun, B.; Chuva de Sousa Lopes, S. M.; Howlett, S. K.; Clarkson, A.; Ahrlund-Richter, L.; Pedersen, R. A.; Vallier, L. *Nature* **2007**, *448*, 191.
16. Tesar, P. J.; Chenoweth, J. G.; Brook, F. A.; Davies, T. J.; Evans, E. P.; Mack, D. L.; Gardner, R. L.; McKay, R. D. G. *Nature* **2007**, *448*, 196.
17. Buehr, M.; Meek, S.; Blair, K.; Yang, J.; Ure, J.; Silva, J.; McLay, R.; Hall, J.; Ying, Q.-L.; Smith, A. *Cell* **2008**, *135*, 1287.
18. Li, P.; Tong, C.; Mehrian-Shai, R.; Jia, L.; Wu, N.; Yan, Y.; Maxson, R. E.; Schulze, E. N.; Song, H.; Hsieh, C.-L.; Pera, M. F.; Ying, Q.-L. *Cell* **2008**, *135*, 1299.
19. Doetschman, T.; Williams, P.; Maeda, N. *Developmental Biology* **1988**, *127*, 224.
20. Graves, K. H.; Moreadith, R. W. *Molecular Reproduction and Development* **1993**, *36*, 424.
21. Stice, S. L.; Strelchenko, N. S.; Keefer, C. L.; Matthews, L. *Biology of Reproduction* **1996**, *54*, 100.
22. Pain, B.; Clark, M. E.; Shen, M.; Nakazawa, H.; Sakurai, M.; Samarut, J.; Etches, R. J. *Development* **1996**, *122*, 2339.
23. Thomson, J. A.; Kalishman, J.; Golos, T. G.; Durning, M.; Harris, C. P.; Becker, R. A.; Hearn, J. P. *Proceedings of the National Academy of Sciences* **1995**, *92*, 7844.
24. Thomson, J. A.; Kalishman, J.; Golos, T. G.; Durning, M.; Harris, C. P.; Hearn, J. P. *Biology of Reproduction* **1996**, *55*, 254.
25. Soriano, P.; Jaenisch, R. *Cell* **1986**, *46*, 19.
26. Resnick, J. L.; Bixler, L. S.; Cheng, L.; Donovan, P. J. *Nature* **1992**, *359*, 550.
27. Matsui, Y.; Zsebo, K.; Hogan, B. L. M. *Cell* **1992**, *70*, 841.
28. Stewart, C. L.; Gadi, I.; Bhatt, H. *Developmental Biology* **1994**, *161*, 626.
29. Rathjen, J.; Lake, J. A.; Bettess, M. D.; Washington, J. M.; Chapman, G.; Rathjen, P. D. *Journal of Cell Science* **1999**, *112*, 601.
30. Chou, Y.-F.; Chen, H.-H.; Eijpe, M.; Yabuuchi, A.; Chenoweth, J. G.; Tesar, P.; Lu, J.; McKay, R. D. G.; Geijsen, N. *Cell* **2008**, *135*, 449.
31. Tanaka, S.; Kunath, T.; Hadjantonakis, A.-K.; Nagy, A.; Rossant, J. *Science* **1998**, *282*, 2072.

32. Niakan, K. K.; Schrode, N.; Cho, L. T. Y.; Hadjantonakis, A.-K. *Nature Protocols* **2013**, *8*, 1028.
33. Narsinh, K. H.; Plews, J.; Wu, J. C. *Molecular Therapy* **2011**, *19*, 635.
34. Nobel Prize in Medicine 2012: http://www.nobelprize.org/nobel_prizes/medicine/laureates/2012
35. Gurdon, J. B.; Melton, D. A. *Science* **2008**, *322*, 1811.
36. Halley-Stott, R. P.; Pasque, V.; Gurdon, J. B. *Development* **2013**, *140*, 2468.
37. Campbell, K. H. S.; McWhir, J.; Ritchie, W. A.; Wilmut, I. *Nature* **1996**, *380*, 64.
38. Wakayama, T.; Perry, A. C. F.; Zuccotti, M.; Johnson, K. R.; Yanagimachi, R. *Nature* **1998**, *394*, 369.
39. Takahashi, K.; Yamanaka, S. *Cell* **2006**, *126*, 663.
40. Phanstiel, D. H.; Brumbaugh, J.; Wenger, C. D.; Tian, S.; Probasco, M. D.; Bailey, D. J.; Swaney, D. L.; Tervo, M. A.; Bolin, J. M.; Ruotti, V.; Stewart, R.; Thomson, J. A.; Coon, J. J. *Nature Methods* **2011**, *8*, 821.
41. Lister, R.; Pelizzola, M.; Kida, Y. S.; Hawkins, R. D.; Nery, J. R.; Hon, G.; Antosiewicz-Bourget, J.; O/'Malley, R.; Castanon, R.; Klugman, S.; Downes, M.; Yu, R.; Stewart, R.; Ren, B.; Thomson, J. A.; Evans, R. M.; Ecker, J. R. *Nature* **2011**, *471*, 68.
42. Gore, A.; Li, Z.; Fung, H.-L.; Young, J. E.; Agarwal, S.; Antosiewicz-Bourget, J.; Canto, I.; Giorgetti, A.; Israel, M. A.; Kiskinis, E.; Lee, J.-H.; Loh, Y.-H.; Manos, P. D.; Montserrat, N.; Panopoulos, A. D.; Ruiz, S.; Wilbert, M. L.; Yu, J.; Kirkness, E. F.; Belmonte, J. C. I.; Rossi, D. J.; Thomson, J. A.; Eggan, K.; Daley, G. Q.; Goldstein, L. S. B.; Zhang, K. *Nature* **2011**, *471*, 63.
43. Hussein, S. M.; Batada, N. N.; Vuoristo, S.; Ching, R. W.; Autio, R.; Narva, E.; Ng, S.; Sourour, M.; Hamalainen, R.; Olsson, C.; Lundin, K.; Mikkola, M.; Trokovic, R.; Peitz, M.; Brustle, O.; Bazett-Jones, D. P.; Alitalo, K.; Lahesmaa, R.; Nagy, A.; Otonkoski, T. *Nature* **2011**, *471*, 58.
44. Zhao, T.; Zhang, Z.-N.; Rong, Z.; Xu, Y. *Nature* **2011**, *474*, 212.
45. Okita, K.; Ichisaka, T.; Yamanaka, S. *Nature* **2007**, *448*, 313.
46. Maherali, N.; Sridharan, R.; Xie, W.; Utikal, J.; Eminli, S.; Arnold, K.; Stadtfeld, M.; Yachechko, R.; Tchieu, J.; Jaenisch, R.; Plath, K.; Hochedlinger, K. *Cell Stem Cell* **2007**, *1*, 55.
47. Wernig, M.; Meissner, A.; Foreman, R.; Brambrink, T.; Ku, M.; Hochedlinger, K.; Bernstein, B. E.; Jaenisch, R. *Nature* **2007**, *448*, 318.
48. Park, I.-H.; Zhao, R.; West, J. A.; Yabuuchi, A.; Huo, H.; Ince, T. A.; Lerou, P. H.; Lensch, M. W.; Daley, G. Q. *Nature* **2008**, *451*, 141.
49. Takahashi, K.; Tanabe, K.; Ohnuki, M.; Narita, M.; Ichisaka, T.; Tomoda, K.; Yamanaka, S. *Cell* **2007**, *131*, 861.
50. Yu, J.; Vodyanik, M. A.; Smuga-Otto, K.; Antosiewicz-Bourget, J.; Frane, J. L.; Tian, S.; Nie, J.; Jonsdottir, G. A.; Ruotti, V.; Stewart, R.; Slukvin, I. I.; Thomson, J. A. *Science* **2007**, *318*, 1917.
51. Buganim, Y.; Markoulaki, S.; van Wietmarschen, N.; Hoke, H.; Wu, T.; Ganz, K.; Akhtar-Zaidi, B.; He, Y.; Abraham, Brian J.; Porubsky, D.; Kulenkampff, E.; Faddah, Dina A.; Shi, L.; Gao, Q.; Sarkar, S.; Cohen, M.; Goldmann, J.; Nery, Joseph R.; Schultz, Matthew D.; Ecker, Joseph R.; Xiao, A.; Young, R. A.; Lansdorp, Peter M.; Jaenisch, R. *Cell Stem Cell* **2014**, *15*, 295.
52. Carey, Bryce W.; Markoulaki, S.; Hanna, Jacob H.; Faddah, Dina A.; Buganim, Y.; Kim, J.; Ganz, K.; Steine, Eveline J.; Cassady, John P.; Creyghton, Menno P.; Welstead, G. G.; Gao, Q.; Jaenisch, R. *Cell Stem Cell* **2011**, *9*, 588.
53. Polo, J. M.; Liu, S.; Figueroa, M. E.; Kulalert, W.; Eminli, S.; Tan, K. Y.; Apostolou, E.; Stadtfeld, M.; Li, Y.; Shioda, T.; Natesan, S.; Wagers, A. J.; Melnick, A.; Evans, T.; Hochedlinger, K. *Nature Biotechnology* **2010**, *28*, 848.
54. Panopoulos, A. D.; Yanes, O.; Ruiz, S.; Kida, Y. S.; Diep, D.; Tautenhahn, R.; Herrerias, A.; Batchelder, E. M.; Plongthongkum, N.; Lutz, M.; Berggren, W. T.; Zhang, K.; Evans, R. M.; Siuzdak, G.; Belmonte, J. C. I. *Cell Research* **2012**, *22*, 168.
55. Zhang, J.; Nuebel, E.; Daley, George Q.; Koehler, Carla M.; Teitell, Michael A. *Cell Stem Cell* **2012**, *11*, 589.
56. Hansson, J.; Rafiee, Mahmoud R.; Reiland, S.; Polo, Jose M.; Gehring, J.; Okawa, S.; Huber, W.; Hochedlinger, K.; Krijgsveld, J. *Cell Reports* **2012**, *2*, 1579.
57. Hussein, S. M. I.; Puri, M. C.; Tonge, P. D.; Benevento, M.; Corso, A. J.; Clancy, J. L.; Mosbergen, R.; Li, M.; Lee, D.-S.; Cloonan, N.; Wood, D. L. A.; Munoz, J.; Middleton, R.; Korn, O.; Patel, H. R.; White, C. A.; Shin, J.-Y.; Gauthier, M. E.; Cao, K.-A. L.; Kim, J.-I.; Mar, J. C.; Shakiba,

N.; Ritchie, W.; Rasko, J. E. J.; Grimmond, S. M.; Zandstra, P. W.; Wells, C. A.; Preiss, T.; Seo, J.-S.; Heck, A. J. R.; Rogers, I. M.; Nagy, A. *Nature* **2014**, *516*, 198.

58. Papp, B.; Plath, K. *Cell* **2013**, *152*, 1324.

59. Yang, C.-S.; Chang, K.-Y.; Rana, Tariq M. *Cell Reports* **2014**, *8*, 327.

60. Qin, H.; Diaz, A.; Blouin, L.; Lebbink, Robert J.; Patena, W.; Tanbun, P.; LeProust, Emily M.; McManus, Michael T.; Song, Jun S.; Ramalho-Santos, M. *Cell* **2014**, *158*, 449.

61. Hu, K. *Stem Cells and Development* **2014**, *23*, 1285.

62. Nakagawa, M.; Takizawa, N.; Narita, M.; Ichisaka, T.; Yamanaka, S. *Proceedings of the National Academy of Sciences* **2010**, *107*, 14152.

63. Nakagawa, M.; Koyanagi, M.; Tanabe, K.; Takahashi, K.; Ichisaka, T.; Aoi, T.; Okita, K.; Mochiduki, Y.; Takizawa, N.; Yamanaka, S. *Nature Biotechnology* **2008**, *26*, 101.

64. Schlaeger, T. M.; Daheron, L.; Brickler, T. R.; Entwisle, S.; Chan, K.; Cianci, A.; DeVine, A.; Ettenger, A.; Fitzgerald, K.; Godfrey, M.; Gupta, D.; McPherson, J.; Malwadkar, P.; Gupta, M.; Bell, B.; Doi, A.; Jung, N.; Li, X.; Lynes, M. S.; Brookes, E.; Cherry, A. B. C.; Demirbas, D.; Tsankov, A. M.; Zon, L. I.; Rubin, L. L.; Feinberg, A. P.; Meissner, A.; Cowan, C. A.; Daley, G. Q. *Nature Biotechnology* **2015**, *33*, 58.

65. Chen, S.; Do, J. T.; Zhang, Q.; Yao, S.; Yan, F.; Peters, E. C.; Schöler, H. R.; Schultz, P. G.; Ding, S. *Proceedings of the National Academy of Sciences* **2006**, *103*, 17266.

66. Ying, Q.-L.; Wray, J.; Nichols, J.; Batlle-Morera, L.; Doble, B.; Woodgett, J.; Cohen, P.; Smith, A. *Nature* **2008**, *453*, 519.

67. Yang, W.; Wei, W.; Shi, C.; Zhu, J.; Ying, W.; Shen, Y.; Ye, X.; Fang, L.; Duo, S.; Che, J.; Shen, H.; Ding, S.; Deng, H. *STEM CELLS* **2009**, *27*, 383.

68. Hanna, J.; Markoulaki, S.; Mitalipova, M.; Cheng, A. W.; Cassady, J. P.; Staerk, J.; Carey, B. W.; Lengner, C. J.; Foreman, R.; Love, J.; Gao, Q.; Kim, J.; Jaenisch, R. *Cell Stem Cell* **2009**, *4*, 513.

69. Kawamata, M.; Ochiya, T. *Proceedings of the National Academy of Sciences* **2010**, *107*, 14223.

70. Li, W.; Wei, W.; Zhu, S.; Zhu, J.; Shi, Y.; Lin, T.; Hao, E.; Hayek, A.; Deng, H.; Ding, S. *Cell Stem Cell* **2009**, *4*, 16.

71. Hanna, J.; Cheng, A. W.; Saha, K.; Kim, J.; Lengner, C. J.; Soldner, F.; Cassady, J. P.; Muffat, J.; Carey, B. W.; Jaenisch, R. *Proceedings of the National Academy of Sciences* **2010**, *107*, 9222.

72. Cao, N.; Liang, H.; Huang, J.; Wang, J.; Chen, Y.; Chen, Z.; Yang, H.-T. *Cell Research* **2013**, *23*, 1119.

73. Watanabe, K.; Ueno, M.; Kamiya, D.; Nishiyama, A.; Matsumura, M.; Wataya, T.; Takahashi, J. B.; Nishikawa, S.; Nishikawa, S.-i.; Muguruma, K.; Sasai, Y. *Nature Biotechnology* **2007**, *25*, 681.

74. Xu, Y.; Zhu, X.; Hahm, H. S.; Wei, W.; Hao, E.; Hayek, A.; Ding, S. *Proceedings of the National Academy of Sciences* **2010**, *107*, 8129.

75. Boitano, A. E.; Wang, J.; Romeo, R.; Bouchez, L. C.; Parker, A. E.; Sutton, S. E.; Walker, J. R.; Flaveny, C. A.; Perdew, G. H.; Denison, M. S.; Schultz, P. G.; Cooke, M. P. *Science* **2010**, *329*, 1345.

76. Chambers, S. M.; Fasano, C. A.; Papapetrou, E. P.; Tomishima, M.; Sadelain, M.; Studer, L. *Nature Biotechnology* **2009**, *27*, 275.

77. Chambers, S. M.; Qi, Y.; Mica, Y.; Lee, G.; Zhang, X.-J.; Niu, L.; Bilsland, J.; Cao, L.; Stevens, E.; Whiting, P.; Shi, S.-H.; Studer, L. *Nature Biotechnology* **2012**, *30*, 715.

78. Borowiak, M.; Maehr, R.; Chen, S.; Chen, A. E.; Tang, W.; Fox, J. L.; Schreiber, S. L.; Melton, D. A. *Cell Stem Cell* **2009**, *4*, 348.

79. Minami, I.; Yamada, K.; Otsuji, Tomomi G.; Yamamoto, T.; Shen, Y.; Otsuka, S.; Kadota, S.; Morone, N.; Barve, M.; Asai, Y.; Tenkova-Heuser, T.; Heuser, John E.; Uesugi, M.; Aiba, K.; Nakatsuji, N. *Cell Reports* **2012**, *2*, 1448.

80. Lian, X.; Hsiao, C.; Wilson, G.; Zhu, K.; Hazeltine, L. B.; Azarin, S. M.; Raval, K. K.; Zhang, J.; Kamp, T. J.; Palecek, S. P. *Proceedings of the National Academy of Sciences* **2012**, *109*, E1848.

81. Kirkeby, A.; Grealish, S.; Wolf, Daniel A.; Nelander, J.; Wood, J.; Lundblad, M.; Lindvall, O.; Parmar, M. *Cell Reports* **2012**, *1*, 703.

82. Mikkelsen, T. S.; Hanna, J.; Zhang, X.; Ku, M.; Wernig, M.; Schorderet, P.; Bernstein, B. E.; Jaenisch, R.; Lander, E. S.; Meissner, A. *Nature* **2008**, *454*, 49.

83. Huangfu, D.; Maehr, R.; Guo, W.; Eijkelenboom, A.; Snitow, M.; Chen, A. E.; Melton, D. A. *Nature Biotechnology* **2008**, *26*, 795.

84. Huangfu, D.; Osafune, K.; Maehr, R.; Guo, W.; Eijkelenboom, A.; Chen, S.; Muhlestein, W.; Melton, D. A. *Nature Biotechnology* **2008**, *26*, 1269.

85. Esteban, M. A.; Wang, T.; Qin, B.; Yang, J.; Qin, D.; Cai, J.; Li, W.; Weng, Z.; Chen, J.; Ni, S.; Chen, K.; Li, Y.; Liu, X.; Xu, J.; Zhang, S.; Li, F.; He, W.; Labuda, K.; Song, Y.; Peterbauer, A.; Wolbank, S.; Redl, H.; Zhong, M.; Cai, D.; Zeng, L.; Pei, D. *Cell Stem Cell* **2010**, *6*, 71.

86. Blaschke, K.; Ebata, K. T.; Karimi, M. M.; Zepeda-Martinez, J. A.; Goyal, P.; Mahapatra, S.; Tam, A.; Laird, D. J.; Hirst, M.; Rao, A.; Lorincz, M. C.; Ramalho-Santos, M. *Nature* **2013**, *500*, 222.

87. Bar-Nur, O.; Brumbaugh, J.; Verheul, C.; Apostolou, E.; Pruteanu-Malinici, I.; Walsh, R. M.; Ramaswamy, S.; Hochedlinger, K. *Nature Methods* **2014**, *11*, 1170.

88. Ichida, J. K.; Blanchard, J.; Lam, K.; Son, E. Y.; Chung, J. E.; Egli, D.; Loh, K. M.; Carter, A. C.; Di Giorgio, F. P.; Koszka, K.; Huangfu, D.; Akutsu, H.; Liu, D. R.; Rubin, L. L.; Eggan, K. *Cell Stem Cell* **2009**, *5*, 491.

89. Maherali, N.; Hochedlinger, K. *Current Biology* **2009**, *19*, 1718.

90. Shi, Y.; Desponts, C.; Do, J. T.; Hahm, H. S.; Schöler, H. R.; Ding, S. *Cell Stem Cell* **2008**, *3*, 568.

91. Yuan, X.; Wan, H.; Zhao, X.; Zhu, S.; Zhou, Q.; Ding, S. *STEM CELLS* **2011**, *29*, 549.

92. Lu, B.; Atala, A. *Drug Discovery Today* **2014**, *19*, 801.

93. Federation, A. J.; Bradner, J. E.; Meissner, A. *Trends in Cell Biology* **2014**, *24*, 179.

94. Hou, P.; Li, Y.; Zhang, X.; Liu, C.; Guan, J.; Li, H.; Zhao, T.; Ye, J.; Yang, W.; Liu, K.; Ge, J.; Xu, J.; Zhang, Q.; Zhao, Y.; Deng, H. *Science* **2013**, *341*, 651.

95. Reymond, J.-L. *Accounts of Chemical Research* **2015**, *48*, 722.

96. Reymond, J.-L.; Awale, M. *ACS Chemical Neuroscience* **2012**, *3*, 649.

97. Reymond, J.-L.; van Deursen, R.; Blum, L. C.; Ruddigkeit, L. *Medicinal Chemistry Communication* **2010**, *1*, 30.

98. Dimmeler, S.; Ding, S.; Rando, T. A.; Trounson, A. *Nature Medicine* **2014**, *20*, 814.

99. Rosemann, A. *Journal of Cellular Biochemistry* **2014**, *115*, 2073.

100. Tabar, V.; Studer, L. *Nature Reviews Genetics* **2014**, *15*, 82.

101. Hirschi, K. K.; Li, S.; Roy, K. *Annual Review of Biomedical Engineering* **2014**, *16*, 277.

102. Volarevic, V.; Nurkovic, J.; Arsenijevic, N.; Stojkovic, M. *STEM CELLS* **2014**, *32*, 2818.

103. Lunn, J. S.; Sakowski, S. A.; Feldman, E. L. *STEM CELLS* **2014**, *32*, 1099.

104. Forbes, S. J.; Rosenthal, N. *Nature Medicine* **2014**, *20*, 857.

105. Ranieri, C.; Claudia Lo, S. In *Scaffolds for Tissue Engineering*; Pan Stanford Publishing, **2014**, pp. 79.

106. Sterneckert, J. L.; Reinhardt, P.; Scholer, H. R. *Nature Reviews Genetics* **2014**, *15*, 625.

107. Grskovic, M.; Javaherian, A.; Strulovici, B.; Daley, G. Q. *Nature Reviews Drug Discovery* **2011**, *10*, 915.

108. Si-Tayeb, K.; Noto, F. K.; Nagaoka, M.; Li, J.; Battle, M. A.; Duris, C.; North, P. E.; Dalton, S.; Duncan, S. A. *Hepatology* **2010**, *51*, 297.

109. Tateishi, K.; He, J.; Taranova, O.; Liang, G.; D'Alessio, A. C.; Zhang, Y. *Journal of Biological Chemistry* **2008**, *283*, 31601.

110. Zwi, L.; Caspi, O.; Arbel, G.; Huber, I.; Gepstein, A.; Park, I.-H.; Gepstein, L. *Circulation* **2009**, *120*, 1513.

111. Karumbayaram, S.; Novitch, B. G.; Patterson, M.; Umbach, J. A.; Richter, L.; Lindgren, A.; Conway, A. E.; Clark, A. T.; Goldman, S. A.; Plath, K.; Wiedau-pazos, M.; Kornblum, H. I.; Lowry, W. E. *STEM CELLS* **2009**, *27*, 806.

112. Ra, J. C.; Shin, I. S.; Kim, S. H.; Kang, S. K.; Kang, B. C.; Lee, H. Y.; Kim, Y. J.; Jo, J. Y.; Yoon, E. J.; Choi, H. J.; Kwon, E. *Stem Cells and Development* **2011**, *20*, 1297.

113. Fan, Y.; Wu, J.; Ashok, P.; Hsiung, M.; Tzanakakis, E. *Stem Cell Reviews and Reports* **2015**, *11*, 96.

114. Jing, D.; Parikh, A.; Canty, J. M.; Tzanakakis, E. S. *Tissue Engineering Part B: Reviews* **2008**, *14*, 393.

115. Liu, Y.; Yang, R.; He, Z.; Gao, W.-Q. *Cell Regeneration* **2013**, *2*, 1.

116. Murphy, S. V.; Atala, A. *BioEssays* **2013**, *35*, 163.

117. Koehler, K. R.; Mikosz, A. M.; Molosh, A. I.; Patel, D.; Hashino, E. *Nature* **2013**, *500*, 217.

118. Takebe, T.; Sekine, K.; Enomura, M.; Koike, H.; Kimura, M.; Ogaeri, T.; Zhang, R.-R.; Ueno, Y.; Zheng, Y.-W.; Koike, N.; Aoyama, S.; Adachi, Y.; Taniguchi, H. *Nature* **2013**, *499*, 481.

119. Nichols, J.; Smith, A. *Cell Stem Cell* **2009**, *4*, 487.
120. Pauklin, S.; Pedersen, R. A.; Vallier, L. *Journal of Cell Science* **2011**, *124*, 3727.
121. Ware, C. B.; Nelson, A. M.; Mecham, B.; Hesson, J.; Zhou, W.; Jonlin, E. C.; Jimenez-Caliani, A. J.; Deng, X.; Cavanaugh, C.; Cook, S.; Tesar, P. J.; Okada, J.; Margaretha, L.; Sperber, H.; Choi, M.; Blau, C. A.; Treuting, P. M.; Hawkins, R. D.; Cirulli, V.; Ruohola-Baker, H. *Proceedings of the National Academy of Sciences* **2014**, *111*, 4484.
122. Theunissen, Thorold W.; Powell, Benjamin E.; Wang, H.; Mitalipova, M.; Faddah, Dina A.; Reddy, J.; Fan, Zi P.; Maetzel, D.; Ganz, K.; Shi, L.; Lungjangwa, T.; Imsoonthornruksa, S.; Stelzer, Y.; Rangarajan, S.; D'Alessio, A.; Zhang, J.; Gao, Q.; Dawlaty, Meelad M.; Young, Richard A.; Gray, Nathanael S.; Jaenisch, R. *Cell Stem Cell* **2014**, *15*, 471.
123. Takashima, Y.; Guo, G.; Loos, R.; Nichols, J.; Ficz, G.; Krueger, F.; Oxley, D.; Santos, F.; Clarke, J.; Mansfield, W.; Reik, W.; Bertone, P.; Smith, A. *Cell* **2014**, *158*, 1254.
124. Cahan, P.; Daley, G. Q. *Nat Reiews Molecular Cell Biolology* **2013**, *14*, 357.
125. Graf, T.; Stadtfeld, M. *Cell Stem Cell* **2008**, *3*, 480.
126. Chambers, I.; Silva, J.; Colby, D.; Nichols, J.; Nijmeijer, B.; Robertson, M.; Vrana, J.; Jones, K.; Grotewold, L.; Smith, A. *Nature* **2007**, *450*, 1230.
127. Singh, A. M.; Hamazaki, T.; Hankowski, K. E.; Terada, N. *STEM CELLS* **2007,** *25*, 2534.
128. Rogers, M. B.; Hosler, B. A.; Gudas, L. J. *Development* **1991**, *113*, 815.
129. Brivanlou, A. H.; Gage, F. H.; Jaenisch, R.; Jessell, T.; Melton, D.; Rossant, J. *Science* **2003**, *300*, 913.
130. Toyooka, Y.; Shimosato, D.; Murakami, K.; Takahashi, K.; Niwa, H. *Development* **2008**, *135*, 909.
131. Chazaud, C.; Yamanaka, Y.; Pawson, T.; Rossant, J. *Developmental Cell* **2006**, *10*, 615.
132. Han, D. W.; Tapia, N.; Joo, J. Y.; Greber, B.; Araúzo-Bravo, M. J.; Bernemann, C.; Ko, K.; Wu, G.; Stehling, M.; Do, J. T.; Schöler, H. R. *Cell* **2010**, *143*, 617.
133. Gardner, R. L.; Lyon, M. F.; Evans, E. P.; Burtenshaw, M. D. *Journal of Embryology and Experimental Morphology* **1985**, *88*, 349.
134. Jang, H. J.; Kim, J. S.; Choi, H. W.; Jeon, I.; Choi, S.; Kim, M. J.; Song, J.; Do, J. T. *Stem Cell Research* **2014**, *12*, 506.
135. Gage, F. H. *Science* **2000**, *287*, 1433.
136. Bao, S.; Tang, F.; Li, X.; Hayashi, K.; Gillich, A.; Lao, K.; Surani, M. A. *Nature* **2009**, *461*, 1292.
137. Yang, J.; van Oosten, A. L.; Theunissen, T. W.; Guo, G.; Silva, J. C. R.; Smith, A. *Cell Stem Cell* **2010**, *7*, 319.
138. Gillich, A.; Bao, S.; Grabole, N.; Hayashi, K.; Trotter, Matthew W. B.; Pasque, V.; Magnúsdóttir, E.; Surani, M. A. *Cell Stem Cell* **2012**, *10*, 425.
139. Guo, G.; Smith, A. *Development* **2010**, *137*, 3185.
140. Guo, G.; Yang, J.; Nichols, J.; Hall, J. S.; Eyres, I.; Mansfield, W.; Smith, A. *Development* **2009**, *136*, 1063.
141. Murayama, H.; Masaki, H.; Sato, H.; Hayama, T.; Yamaguchi, T.; Nakauchi, H. *Stem Cell Reports* **2015**, *4*, 103.
142. Zhou, H.; Li, W.; Zhu, S.; Joo, J. Y.; Do, J. T.; Xiong, W.; Kim, J. B.; Zhang, K.; Schöler, H. R.; Ding, S. *Journal of Biological Chemistry* **2010**, *285*, 29676.
143. Rao, J.; Greber, B. In *Stem Cell Transcriptional Networks;* Kidder, B. L., Ed.; Springer New York, **2014**; Vol. 1150, pp. 215.
144. Bernemann, C.; Greber, B.; Ko, K.; Sterneckert, J.; Han, D. W.; Araúzo-Bravo, M. J.; Schöler, H. R. *STEM CELLS* **2011**, *29*, 1496.
145. Herrmann, B. G.; Labeit, S.; Poustka, A.; King, T. R.; Lehrach, H. *Nature* **1990,** *343*, 617.
146. Papaioannou, V. E. *Development* **2014**, *141*, 3819.
147. Showell, C.; Binder, O.; Conlon, F. L. *Developmental Dynamics* **2004**, *229*, 201.
148. Illich, D. J.; *Discovery and Mechanistic Studies of Triamtereme-Induced Naïve Conversion of Late-Stage Epiblast Stem Cells* **2014** (Doctoral Thesis at Technical University Dortmund).
149. Illich, D.J.; Zhang, M.; Ursu, A.; Osorno, R.; Kim, K.-P.; Yoon, J.; Araúzo-Bravo, M.J.; Wu, G.; Esch, D.; Adachi, K.; Sabour, D.; Grassme, K.S.; Stehling, M.; Höing, S.; Herzog, W.; Ziegler, S.; Chambers, I.; Waldmann, H.; Schöler, H.R. *Distinct signaling requirements for the establishment of ESC pluripotency in late-stage EpiSCs* (submitted).
150. Kellenberger, S.; Gautschi, I.; Schild, L. *Molecular Pharmacology* **2003**, *64*, 848.

151. Silva, J.; Barrandon, O.; Nichols, J.; Kawaguchi, J.; Theunissen, T. W.; Smith, A. *PLoS Biol* **2008**, *6*, e253.
152. Spring, D. R. *Chemical Society Reviews* **2005**, *34*, 472.
153. O' Connor, C. J.; Laraia, L.; Spring, D. R. *Chemical Society Reviews* **2011**, *40*, 4332.
154. Stockwell, B. R. *Nature Reviews Genetics* **2000**, *1*, 116.
155. Schreiber, S. L. *Bioorganic & Medicinal Chemistry* **1998**, *6*, 1127.
156. Kawasumi, M.; Nghiem, P. *Journal of Investigative Dermatology* **2007**, *127*, 1577.
157. Hopkins, A. L.; Groom, C. R. *Nature Reviews Drug Discovery* **2002**, *1*, 727.
158. Haggarty, S. J.; Koeller, K. M.; Wong, J. C.; Grozinger, C. M.; Schreiber, S. L. *Proceedings of the National Academy of Sciences* **2003**, *100*, 4389.
159. Hideshima, T.; Bradner, J. E.; Wong, J.; Chauhan, D.; Richardson, P.; Schreiber, S. L.; Anderson, K. C. *Proceedings of the National Academy of Sciences of the United States of America* **2005**, *102*, 8567.
160. Jin, L.; Wang, W.; Fang, G. *Annual Review of Pharmacology and Toxicology* **2014**, *54*, 435.
161. Milroy, L.-G.; Grossmann, T. N.; Hennig, S.; Brunsveld, L.; Ottmann, C. *Chemical Reviews* **2014**, *114*, 4695.
162. Sander, J. D.; Joung, J. K. *Nature Biotechnology* **2014**, *32*, 347.
163. Hsu, Patrick D.; Lander, Eric S.; Zhang, F. *Cell* **2014**, *157*, 1262.
164. Islam, K. *ACS Chemical Biology* **2015**, *10*, 343.
165. Shokat, K.; Velleca, M. *Drug Discovery Today* **2002**, *7*, 872.
166. Shah, K. *IUBMB Life* **2005**, *57*, 397.
167. Knight, Z. A.; Shokat, K. M. *Cell* **2007**, *128*, 425.
168. Jiang, J.; Wakimoto, H.; Seidman, J. G.; Seidman, C. E. *Science* **2013**, *342*, 111.
169. Hohjoh, H. *Pharmaceuticals* **2013**, *6*, 522.
170. Lee, D.-F.; Su, J.; Ang, Y.-S.; Carvajal-Vergara, X.; Mulero-Navarro, S.; Pereira, Carlos F.; Gingold, J.; Wang, H.-L.; Zhao, R.; Sevilla, A.; Darr, H.; Williamson, Andrew J. K.; Chang, B.; Niu, X.; Aguilo, F.; Flores, Elsa R.; Sher, Y.-P.; Hung, M.-C.; Whetton, Anthony D.; Gelb, Bruce D.; Moore, Kateri A.; Snoeck, H.-W.; Ma'ayan, A.; Schaniel, C.; Lemischka, Ihor R. *Cell Stem Cell* **2012**, *11*, 179.
171. Li, Z.; Rana, T. M. *Nature Communications* **2012**, *3*, 1085.
172. Sakurai, K.; Talukdar, I.; Patil, Veena S.; Dang, J.; Li, Z.; Chang, K.-Y.; Lu, C.-C.; Delorme-Walker, V.; DerMardirossian, C.; Anderson, K.; Hanein, D.; Yang, C.-S.; Wu, D.; Liu, Y.; Rana, Tariq M. *Cell Stem Cell* **2014**, *14*, 523.
173. Ziegler, S.; Pries, V.; Hedberg, C.; Waldmann, H. *Angewandte Chemie International Edition* **2013**, *52*, 2744.
174. Futamura, Y.; Muroi, M.; Osada, H. *Molecular BioSystems* **2013**, *9*, 897.
175. Kawatani, M.; Osada, H. *MedChemComm* **2014**, *5*, 277.
176. Ursu, A.; Waldmann, H. *Bioorganic & Medicinal Chemistry Letters* **2015**, 25, 3079.
177. Besnard, J.; Ruda, G. F.; Setola, V.; Abecassis, K.; Rodriguiz, R. M.; Huang, X.-P.; Norval, S.; Sassano, M. F.; Shin, A. I.; Webster, L. A.; Simeons, F. R. C.; Stojanovski, L.; Prat, A.; Seidah, N. G.; Constam, D. B.; Bickerton, G. R.; Read, K. D.; Wetsel, W. C.; Gilbert, I. H.; Roth, B. L.; Hopkins, A. L. *Nature* **2012**, *492*, 215.
178. Li, W.; Li, K.; Wei, W.; Ding, S. *Cell Stem Cell* **2013**, *13*, 270.
179. Davies, S. G.; Kennewell, P. D.; Russell, A. J.; Seden, P. T.; Westwood, R.; Wynne, G. M. *Journal of Medicinal Chemistry* **2015**, *58*, 2863.
180. Palanki, M. S. S.; Dneprovskaia, E.; Doukas, J.; Fine, R. M.; Hood, J.; Kang, X.; Lohse, D.; Martin, M.; Noronha, G.; Soll, R. M.; Wrasidlo, W.; Yee, S.; Zhu, H. *Journal of Medicinal Chemistry* **2007**, *50*, 4279.
181. Jungmann, O.; Pfleiderer, W. *Nucleosides, Nucleotides & Nucleic Acids* **2009**, *28*, 550.
182. Weinstock, J.; Dunoff, R. Y.; Sutton, B.; Trost, B.; Kirkpatrick, J.; Farina, F.; Straub, A. S. *Journal of Medicinal Chemistry* **1968**, *11*, 549.
183. Marchetti, F.; Cano, C.; Curtin, N. J.; Golding, B. T.; Griffin, R. J.; Haggerty, K.; Newell, D. R.; Parsons, R. J.; Payne, S. L.; Wang, L. Z.; Hardcastle, I. R. *Organic & Biomolecular Chemistry* **2010**, *8*, 2397.

184. Weinstock, J.; Graboyes, H.; Jaffe, G.; Pachter, I. J.; Snader, K.; Karash, C. B.; Dunoff, R. Y. *Journal of Medicinal Chemistry* **1968**, *11*, 560.

185. Osdene, T. S.; Russell, P. B.; Rane, L. *Journal of Medicinal Chemistry* **1967**, *10*, 431.

186. Spickett, R. G. W.; Timmis, G. M. *Journal of the Chemical Society (Resumed)* **1954**, 2887.

187. Yeom, Y. I.; Fuhrmann, G.; Ovitt, C. E.; Brehm, A.; Ohbo, K.; Gross, M.; Hubner, K.; Scholer, H. R. *Development* **1996**, *122*, 881.

188. Kabsch, W. *Acta Crystallographica Section D* **2010**, *66*, 125.

189. McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. *Journal of Applied Crystallography* **2007**, *40*, 658.

190. Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. *Acta Crystallographica Section D* **2011**, *67*, 235.

191. Murshudov, G. N.; Skubak, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A. *Acta Crystallographica Section D* **2011**, *67*, 355.

192. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. *Acta Crystallographica Section D* **2010**, *66*, 486.

193. Cox, J.; Mann, M. *Nature Biotechnology* **2008**, *26*, 1367.

194. McCormack, J. J.; Jaffe, J. J. *Journal of Medicinal Chemistry* **1969**, *12*, 662.

195. Dewey, V. C.; Kidder, G. W. *Biochemical Pharmacology* **1974**, *23*, 773.

196. Davis, M. I.; Hunt, J. P.; Herrgard, S.; Ciceri, P.; Wodicka, L. M.; Pallares, G.; Hocker, M.; Treiber, D. K.; Zarrinkar, P. P. *Nature Biotechnology* **2011**, *29*, 1046.

197. Doukas, J.; Wrasidlo, W.; Noronha, G.; Dneprovskaia, E.; Fine, R.; Weis, S.; Hood, J.; DeMaria, A.; Soll, R.; Cheresh, D. *Proceedings of the National Academy of Sciences* **2006**, *103*, 19866.

198. Sinnberg, B. S. a. T. *Molecular Cancer* **2014**, *13*, 231.

199. Knippschild, U.; Krüger, M.; Richter, J.; Xu, P.; García-Reyes, B.; Peifer, C.; Halekotte, J.; Bakulev, V.; Bischof, J. *Frontiers in Oncology* **2014**, *4*, *1*.

200. Rena, G.; Bain, J.; Elliott, M.; Cohen, P. *EMBO Reports* **2004**, *5*, 60.

201. Kinase-Glo® Luminescent Kinase Assay: https://www.promega.de/en/products/cell-signaling/kinase-assays/kinase_glo-luminescent-kinase-assays.

202. Okabe, M. K. a. T. *Assay and Drug Development Technologies* **2004**, *2*, 153.

203. Kashem, M. A.; Nelson, R. M.; Yingling, J. D.; Pullen, S. S.; Prokopowicz, A. S.; Jones, J. W.; Wolak, J. P.; Rogers, G. R.; Morelock, M. M.; Snow, R. J.; Homon, C. A.; Jakes, S. *Journal of Biomolecular Screening* **2007**, *12*, 70.

204. Leahy, J. W.; Buhr, C. A.; Johnson, H. W. B.; Kim, B. G.; Baik, T.; Cannoy, J.; Forsyth, T. P.; Jeong, J. W.; Lee, M. S.; Ma, S.; Noson, K.; Wang, L.; Williams, M.; Nuss, J. M.; Brooks, E.; Foster, P.; Goon, L.; Heald, N.; Holst, C.; Jaeger, C.; Lam, S.; Lougheed, J.; Nguyen, L.; Plonowski, A.; Song, J.; Stout, T.; Wu, X.; Yakes, M. F.; Yu, P.; Zhang, W.; Lamb, P.; Raeber, O. *Journal of Medicinal Chemistry* **2012**, *55*, 5467.

205. Narsimha Munagala, S. N., Wendy Lam, Jae Lee, Alison Joly, Kirk McMillan, and Wentao Zhang. *ASSAY and Drug Development Technologies* **2007**, *5*, 65.

206. Andrea Baki, A. B., László Molnár, Györgyi Szendrei, and György M. Keserü. *ASSAY and Drug Development Technologies* **2007**, *5*, 75.

207. Salado, I. G.; Redondo, M.; Bello, M. L.; Perez, C.; Liachko, N. F.; Kraemer, B. C.; Miguel, L.; Lecourtois, M.; Gil, C.; Martinez, A.; Perez, D. I. *Journal of Medicinal Chemistry* **2014**, *57*, 2755.

208. Hirota, T.; Lee, J. W.; Lewis, W. G.; Zhang, E. E.; Breton, G.; Liu, X.; Garcia, M.; Peters, E. C.; Etchegaray, J.-P.; Traver, D.; Schultz, P. G.; Kay, S. A. *PLoS Biol* **2010**, *8*, e1000559.

209. Mente, S.; Arnold, E.; Butler, T.; Chakrapani, S.; Chandrasekaran, R.; Cherry, K.; DiRico, K.; Doran, A.; Fisher, K.; Galatsis, P.; Green, M.; Hayward, M.; Humphrey, J.; Knafels, J.; Li, J.; Liu, S.; Marconi, M.; McDonald, S.; Ohren, J.; Paradis, V.; Sneed, B.; Walton, K.; Wager, T. *Journal of Medicinal Chemistry* **2013**, *56*, 6819.

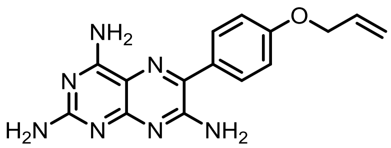
210. Rivers, A.; Gietzen, K. F.; Vielhaber, E.; Virshup, D. M. *Journal of Biological Chemistry* **1998**, *273*, 15980.

211. Cegielska, A.; Gietzen, K. F.; Rivers, A.; Virshup, D. M. *Journal of Biological Chemistry* **1998**, *273*, 1357.
212. Graves, P. R.; Roach, P. J. *Journal of Biological Chemistry* **1995**, *270*, 21689.
213. Zhang, J.-H.; Chung, T. D. Y.; Oldenburg, K. R. *Journal of Biomolecular Screening* **1999**, *4*, 67.
214. Walton, K. M.; Fisher, K.; Rubitski, D.; Marconi, M.; Meng, Q.-J.; Sládek, M.; Adams, J.; Bass, M.; Chandrasekaran, R.; Butler, T.; Griffor, M.; Rajamohan, F.; Serpa, M.; Chen, Y.; Claffey, M.; Hastings, M.; Loudon, A.; Maywood, E.; Ohren, J.; Doran, A.; Wager, T. T. *Journal of Pharmacology and Experimental Therapeutics* **2009**, *330*, 430.
215. https://clinicaltrials.gov/ct2/show/NCT00103350.
216. Long, A. M.; Zhao, H.; Huang, X. *Journal of Medicinal Chemistry* **2012**, *55*, 10307.
217. Long, A.; Zhao, H.; Huang, X. *Journal of Medicinal Chemistry* **2012**, *55*, 956.
218. Simon, G. M.; Niphakis, M. J.; Cravatt, B. F. *Nature Chemical Biology* **2013**, *9*, 200.
219. https://www.thermofisher.com/order/catalog/product/88310.
220. Xiao, Y.; Wang, Y. *Mass Spectrometry Reviews* **2014** (ahead of print).
221. Patricelli, Matthew P.; Nomanbhoy, Tyzoon K.; Wu, J.; Brown, H.; Zhou, D.; Zhang, J.; Jagannathan, S.; Aban, A.; Okerberg, E.; Herring, C.; Nordin, B.; Weissig, H.; Yang, Q.; Lee, J.-D.; Gray, Nathanael S.; Kozarich, John W. *Chemistry & Biology* **2011**, *18*, 699.
222. McAllister, F. E.; Niepel, M.; Haas, W.; Huttlin, E.; Sorger, P. K.; Gygi, S. P. *Analytical Chemistry* **2013**, *85*, 4666.
223. Adachi, J.; Kishida, M.; Watanabe, S.; Hashimoto, Y.; Fukamizu, K.; Tomonaga, T. *Journal of Proteome Research* **2014**, *13*, 5461.
224. Volonté, C.; D'Ambrosi, N. *FEBS Journal* **2009**, *276*, 318.
225. dos Santos-Rodrigues, A.; Grañé-Boladeras, N.; Bicket, A.; Coe, I. R. *Neurochemistry International* **2014**, *73*, 229.
226. Mark Knapp, C. B., Jeremy M. Murray and Dirksen E. Bussiere. *Current Topics in Medicinal Chemistry* **2006**, *6*, 1129.
227. Murray, J.; Bussiere, D. *Chemogenomics;* Jacoby, E., Ed.; Humana Press, **2009**; Vol. 575, pp. 47.
228. Rosenblum, J. S.; Nomanbhoy, T. K.; Kozarich, J. W. *FEBS Letters* **2013**, *587*, 1870.
229. Saraste, M.; Sibbald, P. R.; Wittinghofer, A. *Trends in Biochemical Sciences* **1990**, *15*, 430.
230. Deyrup, A. T.; Krishnan, S.; Cockburn, B. N.; Schwartz, N. B. *Journal of Biological Chemistry* **1998**, *273*, 9450.
231. Sadler, N. C.; Angel, T. E.; Lewis, M. P.; Pederson, L. M.; Chauvigné-Hines, L. M.; Wiedner, S. D.; Zink, E. M.; Smith, R. D.; Wright, A. T. *PLoS ONE* **2012**, *7*, e47996.
232. Ansong, C.; Ortega, C.; Payne, Samuel H.; Haft, Daniel H.; Chauvignè-Hines, Lacie M.; Lewis, Michael P.; Ollodart, Anja R.; Purvine, Samuel O.; Shukla, Anil K.; Fortuin, S.; Smith, Richard D.; Adkins, Joshua N.; Grundner, C.; Wright, Aaron T. *Chemistry & Biology* **2013**, *20*, 123.
233. Ortega, C.; Liao, R.; Anderson, L. N.; Rustad, T.; Ollodart, A. R.; Wright, A. T.; Sherman, D. R.; Grundner, C. *PLoS Biol* **2014**, *12*, e1001746.
234. Okerberg, E.; Shih, A.; Brown, H.; Alemayehu, S.; Aban, A.; Nomanbhoy, T. K. *Current Protocols in Chemical Biology* **2013**, *5*, 213.
235. Edgar, A.J.; Trost, M.; Watts, C.; Zaru, S. *Bioscience Reports* **2014**, *34*, 29.
236. Cravatt, B. F.; Wright, A. T.; Kozarich, J. W. *Annual Review of Biochemistry* **2008**, *77*, 383.
237. Mellacheruvu, D.; Wright, Z.; Couzens, A. L.; Lambert, J.-P.; St-Denis, N. A.; Li, T.; Miteva, Y. V.; Hauri, S.; Sardiu, M. E.; Low, T. Y.; Halim, V. A.; Bagshaw, R. D.; Hubner, N. C.; al-Hakim, A.; Bouchard, A.; Faubert, D.; Fermin, D.; Dunham, W. H.; Goudreault, M.; Lin, Z.-Y.; Badillo, B. G.; Pawson, T.; Durocher, D.; Coulombe, B.; Aebersold, R.; Superti-Furga, G.; Colinge, J.; Heck, A. J. R.; Choi, H.; Gstaiger, M.; Mohammed, S.; Cristea, I. M.; Bennett, K. L.; Washburn, M. P.; Raught, B.; Ewing, R. M.; Gingras, A.-C.; Nesvizhskii, A. I. *Nature Methods* **2013**, *10*, 730.
238. Franceschini, A.; Szklarczyk, D.; Frankild, S.; Kuhn, M.; Simonovic, M.; Roth, A.; Lin, J.; Minguez, P.; Bork, P.; von Mering, C.; Jensen, L. J. *Nucleic Acids Research* **2013**, *41*, D808.
239. Walter, S. A.; Cutler, R. E.; Martinez, R.; Gishizky, M.; Hill, R. J. *Journal of Biological Chemistry* **2003**, *278*, 18221.
240. CellTiter-Glo® Luminescent Cell Viability Assay: https://www.promega.de/en/products/cell-health-and-metabolism/cell-viability-assays/celltiter_glo-luminescent-cell-viability-assay.

241. Roberts, D.; Hall, T. C. *The Journal of Clinical Pharmacology and The Journal of New Drugs* **1968**, *8*, 217.
242. Sidhom, M. B.; Velez, M. R. *Journal of Pharmaceutical and Biomedical Analysis* **1989**, *7*, 1551.
243. Aronov, A. M.; McClain, B.; Moody, C. S.; Murcko, M. A. *Journal of Medicinal Chemistry* **2008**, *51*, 1214.
244. Stegmaier, K. *The Journal of Experimental Medicine* **2014**, *211*, 594.
245. Järås, M.; Miller, P. G.; Chu, L. P.; Puram, R. V.; Fink, E. C.; Schneider, R. K.; Al-Shahrour, F.; Peña, P.; Breyfogle, L. J.; Hartwell, K. A.; McConkey, M. E.; Cowley, G. S.; Root, D. E.; Kharas, M. G.; Mullally, A.; Ebert, B. L. *The Journal of Experimental Medicine* **2014**, *211*, 605.
246. Tiong, K.-L.; Chang, K.-C.; Yeh, K.-T.; Liu, T.-Y.; Wu, J.-H.; Hsieh, P.-H.; Lin, S.-H.; Lai, W.-Y.; Hsu, Y.-C.; Chen, J.-Y.; Chang, J.-G.; Shieh, G. S. *Neoplasia* **2014**, *16*, 441.
247. Stoter, M.; Bamberger, A.-M.; Aslan, B.; Kurth, M.; Speidel, D.; Loning, T.; Frank, H.-G.; Kaufmann, P.; Lohler, J.; Henne-Bruns, D.; Deppert, W.; Knippschild, U. *Oncogene* **2005**, *24*, 7964.
248. Hirner, H.; Günes, C.; Bischof, J.; Wolff, S.; Grothey, A.; Kühl, M.; Oswald, F.; Wegwitz, F.; Bösl, M. R.; Trauzold, A.; Henne-Bruns, D.; Peifer, C.; Leithäuser, F.; Deppert, W.; Knippschild, U. *PLoS ONE* **2012**, *7*, e29709.
249. Bibian, M.; Rahaim, R. J.; Choi, J. Y.; Noguchi, Y.; Schürer, S.; Chen, W.; Nakanishi, S.; Licht, K.; Rosenberg, L. H.; Li, L.; Feng, Y.; Cameron, M. D.; Duckett, D. R.; Cleveland, J. L.; Roush, W. R. *Bioorganic & Medicinal Chemistry Letters* **2013**, *23*, 4374.
250. Sánchez Alvarado, A.; Yamanaka, S. *Cell* **2014**, *157*, 110.
251. Lakshmipathy, U.; Verfaillie, C. *Blood Reviews* **2005**, *19*, 29.
252. Kuramochi, S.; Moriguchi, T.; Kuida, K.; Endo, J.; Semba, K.; Nishida, E.; Karasuyama, H. *Journal of Biological Chemistry* **1997**, *272*, 22679.
253. Endo, J.; Toyama-Sorimachi, N.; Taya, C.; Kuramochi-Miyagawa, S.; Nagata, K.; Kuida, K.; Takashi, T.; Yonekawa, H.; Yoshizawa, Y.; Miyasaka, N.; Karasuyama, H. *FEBS Letters* **2000**, *468*, 234.
254. Belkina, N. V.; Liu, Y.; Hao, J.-J.; Karasuyama, H.; Shaw, S. *Proceedings of the National Academy of Sciences* **2009**, *106*, 4707.
255. Fukumura, K., Yamashita, Y., Kawazu, M., Sai, E., Fujiwara, S., Nakamura, N., Takeuchi, K., Ando, M., Miyazono, K., Ueno, T., Ozawa, K., Mano, H. *Oncology Reports* **2013**, *30*, 1542.
256. Shilpi Arora, I. M. G., R Tanner Hagelstrom, Christian Beaudry, Ashish Choudhary, Chao Sima, Raoul Tibes, Spyro Mousses and David O Azorsa. *Molecular Cancer* **2010**, *9*, 218.
257. Fiume, R.; Stijf-Bultsma, Y.; Shah, Z. H.; Keune, W. J.; Jones, D. R.; Jude, J. G.; Divecha, N. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids* **2015**, *1851*, 898.
258. Emerling, Brooke M.; Hurov, Jonathan B.; Poulogiannis, G.; Tsukazawa, Kazumi S.; Choo-Wing, R.; Wulf, Gerburg M.; Bell, Eric L.; Shim, H.-S.; Lamia, Katja A.; Rameh, Lucia E.; Bellinger, G.; Sasaki, Atsuo T.; Asara, John M.; Yuan, X.; Bullock, A.; DeNicola, Gina M.; Song, J.; Brown, V.; Signoretti, S.; Cantley, Lewis C. *Cell* **2013**, *155*, 844.
259. Bultsma, Y.; Keune, W.-J.; Divecha, N. *Biochemical Journal* **2010**, *430*, 223.
260. Wang, M.; Bond, Nicholas J.; Letcher, Andrew J.; Richardson, Jonathan P.; Lilley, Kathryn S.; Irvine, Robin F.; Clarke, Jonathan H. *Biochemical Journal* **2010**, *430*, 215.
261. Jude, J. G.; Spencer, G. J.; Huang, X.; Somerville, T. D. D.; Jones, D. R.; Divecha, N.; Somervaille, T. C. P. *Oncogene* **2015**, *34*, 1253.
262. Ellinger-Ziegelbauer, H.; Karasuyama, H.; Yamada, E.; Tsujikawa, K.; Todokoro, K.; Nishida, E. *Genes to Cells* **2000**, *5*, 491.
263. Jackson, S. P. *The International Journal of Biochemistry & Cell Biology* **1997**, *29*, 935.
264. Hunter, T. *Cell* **1995**, *83*, 1.
265. Hartley, K. O.; Gell, D.; Smith, G. C. M.; Zhang, H.; Divecha, N.; Connelly, M. A.; Admon, A.; Lees-Miller, S. P.; Anderson, C. W.; Jackson, S. P. *Cell* **1995**, *82*, 849.
266. Abraham, R. T. *DNA Repair* **2004**, *3*, 883.
267. Carreras Puigvert, J.; von Stechow, L.; Siddappa, R.; Pines, A.; Bahjat, M.; Haazen, L. C. J. M.; Olsen, J. V.; Vrieling, H.; Meerman, J. H. N.; Mullenders, L. H. F.; van de Water, B.; Danen, E. H. J. *Science Signaling* **2013**, *6*, 1.
268. Yao, K.; Ki, M. O.; Chen, H.; Cho, Y.-Y.; Kim, S.-H.; Yu, D. H.; Lee, S.-Y.; Lee, K.-Y.; Bae, K.; Peng, C.; Lim, D. Y.; Bode, A. M.; Dong, Z. *Stem Cell Research* **2014**, *12*, 139.
269. Wilson, J. S. K. a. C. H. *Biochemical Society Transactions* **2013**, *41*, 1055.

270. Zeqiraj, E.; van Aalten, D. M. F. *Current Opinion in Structural Biology* **2010**, *20*, 772.
271. Boudeau, J.; Miranda-Saavedra, D.; Barton, G. J.; Alessi, D. R. *Trends in Cell Biology* **2006**, *16*, 443.
272. Wilson, C. H.; Crombie, C.; van der Weyden, L.; Poulogiannis, G.; Rust, A. G.; Pardo, M.; Gracia, T.; Yu, L.; Choudhary, J.; Poulin, G. B.; McIntyre, R. E.; Winton, D. J.; March, H. N.; Arends, M. J.; Fraser, A. G.; Adams, D. J. *The EMBO Journal* **2012**, *31*, 2486.
273. van den Berg, D. L. C.; Snoek, T.; Mullin, N. P.; Yates, A.; Bezstarosti, K.; Demmers, J.; Chambers, I.; Poot, R. A. *Cell Stem Cell* **2010**, *6*, 369.
274. Yang, J.; Chai, L.; Fowles, T. C.; Alipio, Z.; Xu, D.; Fink, L. M.; Ward, D. C.; Ma, Y. *Proceedings of the National Academy of Sciences* **2008**, *105*, 19756.
275. Tan, E. H.; Sansom, O. J. *The EMBO Journal* **2012**, *31*, 2444.
276. Ruiz, C.; Oeggerli, M.; Germann, M.; Gluderer, S.; Stocker, H.; Andreozzi, M.; Thalmann, G. N.; Cecchini, M. G.; Zellweger, T.; Stürm, S.; Koivisto, P. A.; Helin, H. J.; Gelmann, E. P.; Glass, A. G.; Gasser, T. C.; Terracciano, L. M.; Bachmann, A.; Wyler, S.; Bubendorf, L.; Rentsch, C. A. *The Prostate* **2012**, *72*, 1678.
277. Tran, K. A.; Jackson, S. A.; Olufs, Z. P. G.; Zaidan, N. Z.; Leng, N.; Kendziorski, C.; Roy, S.; Sridharan, R. *Nature Communications* **2015**, *6*, 6188.
278. Alsheich-Bartok, O.; Haupt, S.; Alkalay-Snir, I.; Saito, S.; Appella, E.; Haupt, Y. *Oncogene* **2008**, *27*, 3653.
279. Dumaz, N.; Milne, D. M.; Meek, D. W. *FEBS Letters* **1999**, *463*, 312.
280. Venerando, A.; Marin, O.; Cozza, G.; Bustos, V.; Sarno, S.; Pinna, L. *Cellular and Molecular Life Sciences* **2010**, *67*, 1105.

11. Compound characterization



6-(4-(allyloxy)phenyl)pteridine-2,4,7-triamine

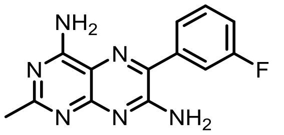
Yield = 56 %

^{1}H-NMR [400 MHz, DMSO-d6] δ: 7.67 (d, *J* = 8.7 Hz, 2H), 7.13 – 6.87 (m, 4H), 6.58 (s, 2H), 6.17 (s, 2H), 6.08 (s, 1H), 5.43 (d, *J* = 15.7 Hz, 1H), 5.28 (d, *J* = 11.9 Hz, 1H), 4.63 (d, *J* = 5.1 Hz, 2H).

^{13}C NMR [101 MHz, DMSO-d6] δ: 162.88, 161.99, 158.42, 155.84, 155.48, 137.01, 135.90, 129.73, 129.71, 128.44, 127.83, 127.60, 114.82, 113.74, 69.22.

FT-IR (cm^{-1}): 3383, 3111, 1610, 1566, 1536, 1425, 1348, 1246, 1177, 840, 817, 778, 754, 728, 684.

HRMS Calcd for C15H16N7O 310.14108 [M+H$^+$]; Found: 310.14145



6-(3-fluorophenyl)-2-methylpteridine-4,7-diamine **(entry 38)**

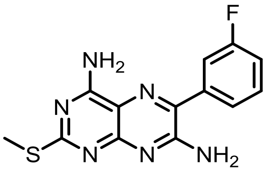
Yield = 52 %

^{1}H-NMR [500 MHz, DMSO-d6] δ: 7.63 (d, J = 7.8 Hz, 2H), 7.55 (d, J = 6.1 Hz, 1H), 7.41 (s, 2H), 7.32 (dd, J = 9.8, 7.2 Hz, 1H), 7.05 (s, 2H), 2.37 (s, 3H).

^{13}C NMR [101 MHz, DMSO-d6] δ: 167.56, , 161.56, 160.98 (d, J = 244.48 Hz, C-F), 155.82, 153.94, 139.99 (d, J = 2.5 Hz), 138.71 (d, J = 7.8 Hz), 130.61 (d, J = 8.4 Hz), 124.51 (d, J = 2.7 Hz), 116.05, 115.79 (d, J = 11.0 Hz), 115.47 (d, J = 7.3 Hz), 26.24.

FT-IR (cm^{-1}): 3465, 3281, 3043, 1633, 1540, 1411, 1332.

HRMS Calcd for C13H12FN6 271.11020 [M+H$^+$]; Found: 271.11042



6-(3-fluorophenyl)-2-(methylthio)pteridine-4,7-diamine **(entry 40)**

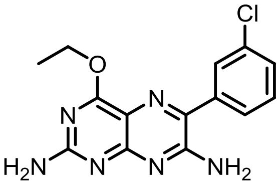
Yield = 47 %

^{1}H-NMR [400 MHz, DMSO-d6] δ: 7.59 (m, 5H), 7.32 (t, J = 8.2 Hz, 1H), 7.17 (s, 2H), 2.46 (s, 3H).

^{13}**C-NMR** [101 MHz, DMSO-d6] δ: 170.86, 162.20 (d, J = 244.42 Hz, C-F), 160.74, 155.93, 153.71, 139.02, 138.63 (d, J = 7.9 Hz), 130.60 (d, J = 8.4 Hz), 124.47 (d, J = 2.7 Hz), 116.01, 115.75 (d, J = 10.4 Hz), 115.40 (d, J = 15.1 Hz), 13.62.

FT-IR (cm^{-1}): 3494, 3282, 3044, 1626, 1531, 1336, 435.

HRMS Calcd for C13H12FN6S 303.08227 [M+H^{+}]; Found: 303.08255.



6-(3-chlorophenyl)-4-ethoxypteridine-2,7-diamine (**entry 50**)

Yield = 44 %

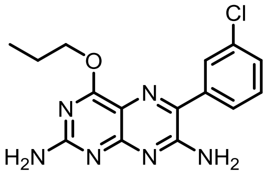
R$_{f}$ = 0.12 [Acetone : PE = 4 : 1]

^{1}**H-NMR** [400 MHz, DMSO-d6] δ: 7.62 – 7.49 (m, 4H), 6.92 (s, 2H), 6.78 (s, 2H), 4.45 (q, J = 7.0 Hz, 2H), 1.36 (t, J = 7.1 Hz, 3H).

^{13}**C-NMR** [101 MHz, DMSO-d6] δ: 165.84, 161.92, 157.63, 155.76, 138.99, 137.30, 133.21, 130.48, 128.51, 128.24, 127.08, 114.19, 62.05, 14.38.

FT-IR (cm^{-1}): 3320, 3202, 3055, 2988, 1630, 1532, 1407, 1315, 1255, 1196, 1099, 1016, 816, 763.

HRMS Calcd for C14H14^{35}ClN6O 317.09121 [M+H^{+}] and C14H14^{37}ClN6O 319.08826 [M+H^{+}]; found: 317.09174 and 319.08850.



6-(3-chlorophenyl)-4-propoxypteridine-2,7-diamine **(entry 51)**

Yield = 92 %

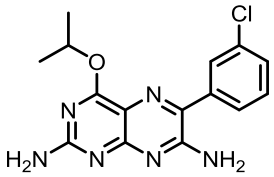
R$_{f}$ = 0.14 [Acetone : PE = 4 : 1]

^{1}**H-NMR** [400 MHz, DMSO-d6] δ: 7.64 – 7.51 (m, 4H), 6.94 (s, 2H), 6.80 (s, 2H), 4.36 (t, *J* = 6.9 Hz, 2H), 1.79 (dt, *J* = 14.4, 7.2 Hz, 2H), 0.96 (t, *J* = 7.4 Hz, 3H);

^{13}**C-NMR** [101 MHz, DMSO-d6] δ: 166.00, 161.90, 157.61, 155.78, 139.01, 137.30, 133.24, 130.51, 128.53, 128.24, 127.09, 114.21, 67.65, 21.67, 10.27;

FT-IR (cm^{-1}): 3328, 3121, 2970, 1629, 1583, 1565, 1534, 1487, 1429, 1406, 1318, 1255, 1195, 1113, 1019, 819, 763, 692.

HRMS Calcd for C15H16^{35}ClN6O 331.10686 [M+H^{+}] and C15H16^{37}ClN6O 333.10391 [M+H^{+}]; Found: 331.10771 and 333.10432.



6-(3-chlorophenyl)-4-isopropoxypteridine-2,7-diamine **(entry 52)**

Yield = 77 %

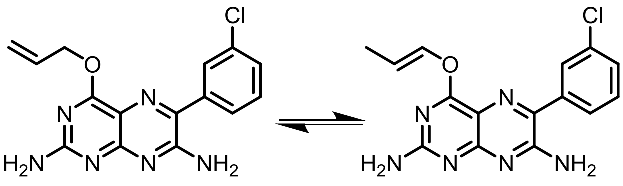
R_f = 0.13 [Acetone : PE = 4 : 1]

^{1}H-NMR [400 MHz, DMSO-d6] δ: 7.62 – 7.49 (m, 4H), 6.89 (s, 2H), 6.74 (s, 2H), 5.45 (dt, J = 12.7, 6.3 Hz, 1H), 1.34 (d, J = 6.2 Hz, 6H).

^{13}C-NMR [101 MHz, DMSO-d6] δ: 165.43, 161.94, 157.73, 155.71, 139.01, 137.28, 133.21, 130.49, 128.49, 128.28, 127.15, 114.36, 68.87, 21.74.

FT-IR (cm^{-1}): 3306, 3153, 2982, 1585, 1562, 1536, 1420, 1313, 1246, 1105, 817, 758, 684, 606.

HRMS Calcd for C15H16^{35}ClN6O 331.10686 [M+H$^+$] and C15H16^{37}ClN6O 333.10391 [M+H$^+$]; Found: 331.10788 and 333.10457.



1 **2**

4-(allyloxy)-6-(3-chlorophenyl)pteridine-2,7-diamine (**entry 53**)

Yield = 76 %

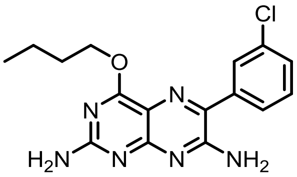
R_f = 0.11 [Acetone : PE = 4 : 1]

^{1}H NMR [400 MHz, DMSO-d6] δ: 7.73 – 7.43 (m, 4H), 6.97 (s, 2H), 6.82 (s, 2H), 6.12 (ddd, J = 16.1, 10.9, 5.7 Hz, 1H), 5.42 (dd, J = 17.2, 1.7 Hz, 1H), 5.26 (d, J = 1.7 Hz, 1H), 4.94 (d, J = 5.7 Hz, 2H), *4.08 (q, J = 5.2 Hz, 1H), 3.17 (d, J = 5.2 Hz, 3H).*

^{13}C-NMR [101 MHz, DMSO-d6] δ: 165.44, 161.79, 157.72, 155.83, 138.95, 137.41, 133.22, 133.09, 130.48, 128.53, 128.20, 127.05, 118.43, 114.07, 66.55, *48.56*.

FT-IR (cm^{-1}): 3312, 3138, 2245, 1578, 1561, 1535, 1485, 1412, 1349, 1316, 1113, 817, 757, 574. **HRMS** Calcd for C15H14^{35}ClN6O 329.09121 [M+H$^+$] and C15H14^{37}ClN6O 331.08826 [M+H$^+$]; Found: 329.09220 and 331.08892.

(Signals in Italic suggest the presence of the structure **2**)



4-butoxy-6-(3-chlorophenyl)pteridine-2,7-diamine **(entry 54)**

Yield = 92 %

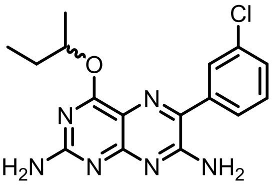
R_f = 0.15 [Acetone : PE = 4 : 1]

^{1}H-NMR [400 MHz, DMSO-d6] δ: 7.66 – 7.48 (m, 4H), 6.92 (s, 2H), 6.78 (s, 2H), 4.40 (t, J = 6.8 Hz, 2H), 1.81 – 1.68 (m, 2H), 1.48 – 1.33 (m, 2H), 0.93 (t, J = 7.4 Hz, 3H).

^{13}C-NMR [101 MHz, DMSO-d6] δ: 165.98, 161.90, 157.62, 155.77, 139.01, 137.25, 133.22, 130.49, 128.50, 128.23, 127.07, 114.21, 112.82, 65.93, 30.35, 19.52, 18.67, 13.66.

FT-IR (cm^{-1}): 3303, 3118, 2959, 2872, 1536, 1488, 1468, 1427, 1349, 1322, 1252, 1203, 1112, 1093, 817, 776, 641, 594.

HRMS Calcd for C16H18^{35}ClN6O 345.12251 [M+H^{+}] and C16H18^{37}ClN6O 347.11956 [M+H^{+}]; Found: 345.12309 and 347.11991.



4-(sec-butoxy)-6-(3-chlorophenyl)pteridine-2,7-diamine **(entry 55)**

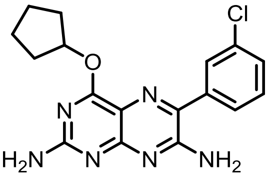
Yield = 64 %

R_f = 0.1 [Acetone : PE = 4 : 1]

^{1}H-NMR [400 MHz, DMSO-d6] δ: 7.62 – 7.50 (m, 4H), 6.89 (s, 2H), 6.74 (s, 2H), 5.30 (h, *J* = 6.1 Hz, 1H), 1.78 – 1.61 (m, 2H), 1.31 (d, *J* = 6.2 Hz, 3H), 0.91 (t, *J* = 7.4 Hz, 3H); **^{13}C-NMR** [101 MHz, DMSO-d6] δ: 165.75, 161.93, 157.72, 155.74, 139.03, 137.26, 133.23, 130.51, 128.51, 128.28, 127.16, 114.41, 73.38, 28.43, 19.21, 9.73;

FT-IR (cm^{-1}): 3311, 3178, 2972, 2935, 2876, 1621, 1584, 1561, 1536, 1424, 1360, 1302, 1246, 1092, 817, 685, 646.

HRMS Calcd for C16H18^{35}ClN6O 345.12251 [M+H^{+}] and C16H18^{37}ClN6O 347.11956 [M+H^{+}]; Found: 345.12375 and 347.12038.



6-(3-chlorophenyl)-4-(cyclopentyloxy)pteridine-2,7-diamine **(entry 56)**

Yield = 31 %

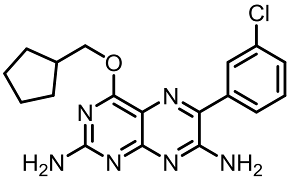
R_f = 0.10 [Acetone : PE = 4 : 1]

^{1}H-NMR [400 MHz, DMSO-d6] δ: 7.62 – 7.49 (m, 4H), 6.90 (s, 2H), 6.76 (s, 2H), 5.51 (tt, *J* = 7.1, 3.4 Hz, 1H), 2.07 – 1.95 (m, 2H), 1.79 –1.67 (m, 4H), 1.69 – 1.51 (m, 2H);

^{13}C-NMR [101 MHz, DMSO-d6] δ: 165.70, 161.86, 157.57, 155.72, 139.02, 137.29, 133.24, 130.52, 128.51, 128.26, 127.15, 114.43, 78.34, 32.22, 23.46;

FT-IR (cm^{-1}): 3320, 3180, 2961, 2872, 1566, 1535, 1468, 1445, 1423, 1408, 1325, 1254, 1190, 815, 772, 756, 690.

HRMS Calcd for $C17H18^{35}ClN6O$ 357.12251 [$M+H^+$] and $C17H18^{37}ClN6O$ 359.11956 [$M+H^+$]; Found: 357.12377 and 359.12043.



6-(3-chlorophenyl)-4-(cyclopentylmethoxy)pteridine-2,7-diamine (**entry 57**)

Yield = 44 %

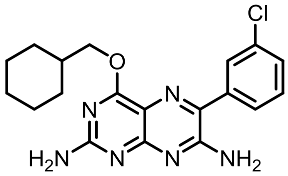
R_f = 0.10 [Acetone : PE = 4 : 1]

^{1}H-NMR [400 MHz, DMSO-d6] δ: 7.67 – 7.51 (m, 4H), 6.94 (s, 2H), 6.79 (s, 2H), 4.30 (d, *J* = 7.2 Hz, 2H), 2.40 (dt, *J* = 11.7, 7.7 Hz, 1H), 1.81 – 1.70 (m, 2H), 1.67 – 1.49 (m, 4H), 1.39 – 1.27 (m, 2H);

^{13}C-NMR [101 MHz, DMSO-d6] δ: 166.11, 161.89, 157.60, 155.76, 139.03, 137.21, 133.24, 130.51, 128.50, 128.23, 127.06, 114.22, 70.02, 38.14, 29.06, 24.80;

FT-IR (cm^{-1}): 3325, 3153, 2951, 2868, 2154, 1580, 1535, 1483, 1424, 1353, 1318, 1253, 1193, 1095, 772, 685, 575.

HRMS Calcd for $C18H20^{35}ClN6O$ 371.13816 [$M+H^+$] and $C18H20^{37}ClN6O$ 373.13521 [$M+H^+$]; found: 371.13920 and 373.13590.



6-(3-chlorophenyl)-4-(cyclohexylmethoxy)pteridine-2,7-diamine (**entry 58**)

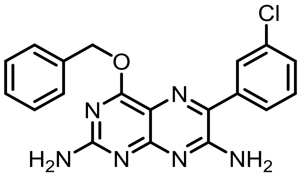
Yield = 85 %

R_f = 0.10 [Acetone : PE = 4 : 1]

^{1}H-NMR [400 MHz, DMSO-d6] δ: 7.70 – 7.43 (m, 4H), 6.91 (s, 2H), 6.77 (s, 2H), 4.21 (d, J = 6.6 Hz, 2H), 1.88 – 1.61 (m, 6H), 1.31 – 1.00 (m, 5H).

^{13}C-NMR [101 MHz, DMSO-d6] δ: 166.08, 161.88, 157.64, 155.78, 139.04, 137.27, 133.23, 130.53, 128.51, 128.24, 127.11, 114.23, 71.05, 36.57, 29.19, 25.97, 25.09.

FT-IR (cm^{-1}): 3317, 3126, 2924, 2851, 1618, 1581, 1559, 1536, 1485, 1425, 1358, 1319, 685. **HRMS** Calcd for C19H22^{35}ClN6O 385.15381 [M+H$^+$] and C19H22^{37}ClN6O 387.15192 [M+H$^+$]; found: 385.15503 and 387.15192.



4-(benzyloxy)-6-(3-chlorophenyl)pteridine-2,7-diamine **(entry 59)**

Yield = 65 %

R_f = 0.13 [Acetone : PE = 4 : 1]

^{1}H-NMR [400 MHz, DMSO-d6] δ: 7.60 (t, J = 2.1 Hz, 1H), 7.57 – 7.54 (m, 1H), 7.52 – 7.48 (m, 4H), 7.36 (m, 3H), 6.96 (s, 2H), 6.86 (s, 2H), 5.50 (s, 2H).

^{13}C-NMR [101 MHz, DMSO-d6] δ: 165.62, 161.74, 157.83, 155.85, 138.90, 137.52, 136.46, 133.21, 130.49, 128.53, 128.37, 128.21, 128.03, 127.10, 114.09, 112.83, 67.31.

FT-IR (cm^{-1}): 3318, 3137, 2962, 1606, 1577, 1556, 1532, 1474, 1425, 1349, 1316, 1259, 1184, 1094, 1021, 817, 799, 745, 696, 589.

HRMS Calcd for C19H16^{35}ClN6O 379.10686 [M+H$^+$] and C19H16^{37}ClN6O [M+H$^+$] 381.10391; Found: 379.10795 and 381.10482.

Curriculum Vitae

ANDREI URSU

Place of birth: Bacau, Romania

Nationality: Romanian

E-mail: andrei.ursu@mpi-dortmund.mpg.de

EDUCATION

2008-2010	***Master of Science*** in Chemistry *("very good")* *Faculty of Chemistry*, *Technical University Munich* (TUM), Germany [DAAD scholarship] Thesis: "Design, synthesis and characterization of novel anion receptors"
2004-2008	***Bachelor of Science*** in Chemistry *(First-Class Honours*), *Faculty of Chemistry*, *University of Bucharest*, Romania
2007-2008	***Internship*** within Department of Inorganic Chemistry, *Faculty of Chemistry*, *University of Bucharest*, Romania
2007	***Erasmus Scholarship*** at the Faculty of Chemistry, *University of Regensburg*, Germany
2000-2004	"Gheorghe Vranceanu" National High-School, Bacau, Romania

CONFERENCES-WORKSHOPS-AWARDS

2015	*4th RIKEN-Max Planck Symposium*, Kobe, Japan *(Oral Presentation)*
2015	*10th Status Seminar Chemical Biology*, Frankfurt am Main, Germany *(Poster Presentation)*
2014	Masaryk University, Brno, Czech Republic *(Invited Lecture)*
2014	*"Cell Based Assays for Screening Workshop"* European Screening Port, Hamburg, Germany *(Workshop)*
2014	*"ReMaT - Research Management Training"* TuTech Innovation, Brussels, Belgium *(Workshop)*
2012	*4th IMPRS International Symposium on Chemical Biology*, Dortmund, Germany *(Oral Presentation)*

2011	“*SupraChem-Symposium in Supramolecular Chemistry”,* Essen, Germany *(Poster Presentation)*
2010	*3rd EuCheMS Chemistry Congress,* Nürnberg, Germany *(Oral Presentation)*
2008	*6th International Congress of Young Chemists “YoungChem”*, Krakow, Poland (*Poster Presentation)*
2008	5th *Students for Students International Conference*, Cluj-Napoca, Romania (*1st prize – Poster Presentation)*
2004	*“Diploma of Excellence”* awarded by “G. Vranceanu” National High School for achievements at national competitions
2003-2004	Member of the Romanian Chemistry Team preparing for the International Chemistry Olympiad (IChO)
2000-2004	Over 20 prizes at national chemistry contests