WILEY

TDM: Emerging Solutions

STM FACT Seminar

London, 20 May 2013

Duncan Campbell

equipment, and to Dr. G. E. R. Deacon and the captain and officers of R.R.S. *Discovery II* for their part in making the observations.

[1] Young, F. B., Gerrard, H., and Jevons, W., *Phil. Mag.*, **40**, 149 (1920).
[2] Longuet-Higgins, M. S., *Mon. Not. Roy. Astro. Soc., Geophys. Supp.*, **5**, 285 (1949).
[3] Von Arx, W. S., Woods Hole Papers in Phys. Oceanog. Meteor., **11** (3) (1950).
[4] Ekman, V. W., *Arkiv. Mat. Astron. Fysik. (Stockholm)*, **2** (11) (1905).

MOLECULAR STRUCTURE OF NUCLEIC ACIDS

A Structure for Deoxyribose Nucleic Acid

We wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). This structure has novel features which are of considerable biological interest.

A structure for nucleic acid has already been proposed by Pauling and Corey[1]. They kindly made their manuscript available to us in advance of publication. Their model consists of three intertwined chains, with the phosphates near the fibre axis, and the bases on the outside. In our opinion, this structure is unsatisfactory for two reasons: (1) We believe that the material which gives the X-ray diagrams is the salt, not the free acid. Without the acidic hydrogen atoms it is not clear what forces would hold the structure together, especially as the negatively charged phosphates near the axis will repel each other. (2) Some of the van der Waals distances appear to be too small.

Another three-chain structure has also been suggested by Fraser (in the press). In his model the phosphates are on the outside and the bases on the inside, linked together by hydrogen bonds. This structure as described is rather ill-defined, and for this reason we shall not comment on it.

This figure is purely diagrammatic. The two ribbons symbolize the two phosphate—sugar chains, and the horizontal rods the pairs of bases holding the chains together. The vertical line marks the fibre axis

We wish to put forward a radically different structure for the salt of deoxyribose nucleic acid. This structure has two helical chains each coiled round the same axis (see diagram). We have made the usual chemical assumptions, namely, that each chain consists of phosphate diester groups joining β-D-deoxyribofuranose residues with 3′,5′ linkages. The two chains (but not their bases) are related by a dyad perpendicular to the fibre axis. Both chains follow right-handed helices, but owing to the dyad the sequences of the atoms in the two chains run in opposite directions. Each chain loosely resembles Furberg's[2] model No. 1; that is, the bases are on the inside of the helix and the phosphates on the outside. The configuration of the sugar and the atoms near it is close to Furberg's 'standard configuration', the sugar being roughly perpendicular to the attached base. There is a residue on each chain every 3·4 A. in the *z*-direction. We have assumed an angle of 36° between adjacent residues in the same chain, so that the structure repeats after 10 residues on each chain, that is, after 34 A. The distance of a phosphorus atom from the fibre axis is 10 A. As the phosphates are on the outside, cations have easy access to them.

The structure is an open one, and its water content is rather high. At lower water contents we would expect the bases to tilt so that the structure could become more compact.

The novel feature of the structure is the manner in which the two chains are held together by the purine and pyrimidine bases. The planes of the bases are perpendicular to the fibre axis. They are joined together in pairs, a single base from one chain being hydrogen-bonded to a single base from the other chain, so that the two lie side by side with identical *z*-co-ordinates. One of the pair must be a purine and the other a pyrimidine for bonding to occur. The hydrogen bonds are made as follows: purine position 1 to pyrimidine position 1; purine position 6 to pyrimidine position 6.

If it is assumed that the bases only occur in the structure in the most plausible tautomeric forms (that is, with the keto rather than the enol configurations) it is found that only specific pairs of bases can bond together. These pairs are: adenine (purine) with thymine (pyrimidine), and guanine (purine) with cytosine (pyrimidine).

In other words, if an adenine forms one member of a pair, on either chain, then on these assumptions the other member must be thymine; similarly for guanine and cytosine. The sequence of bases on a single chain does not appear to be restricted in any way. However, if only specific pairs of bases can be formed, it follows that if the sequence of bases on one chain is given, then the sequence on the other chain is automatically determined.

It has been found experimentally[3,4] that the ratio of the amounts of adenine to thymine, and the ratio of guanine to cytosine, are always very close to unity for deoxyribose nucleic acid.

It is probably impossible to build this structure with a ribose sugar in place of the deoxyribose, as the extra oxygen atom would make too close a van der Waals contact.

The previously published X-ray data[5,6] on deoxyribose nucleic acid are insufficient for a rigorous test of our structure. So far as we can tell, it is roughly compatible with the experimental data, but it must be regarded as unproved until it has been checked against more exact results. Some of these are given in the following communications. We were not aware of the details of the results presented there when we devised our structure, which rests mainly though not entirely on published experimental data and stereochemical arguments.

It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.

Full details of the structure, including the conditions assumed in building it, together with a set of co-ordinates for the atoms, will be published elsewhere.

We are much indebted to Dr. Jerry Donohue for constant advice and criticism, especially on interatomic distances. We have also been stimulated by a knowledge of the general nature of the unpublished experimental results and ideas of Dr. M. H. F. Wilkins, Dr. R. E. Franklin and their co-workers at King's College, London. One of us (J. D. W.) has been aided by a fellowship from the National Foundation for Infantile Paralysis.

J. D. Watson
F. H. C. Crick
Medical Research Council Unit for the Study of the Molecular Structure of Biological Systems,
Cavendish Laboratory, Cambridge.
April 2.

[1] Pauling, L., and Corey, R. B., *Nature*, **171**, 346 (1953); *Proc. U.S. Nat. Acad. Sci.*, **39**, 84 (1953).
[2] Furberg, S., *Acta Chem. Scand.*, **6**, 634 (1952).
[3] Chargaff, E., for references see Zamenhof, S., Brawerman, G., and Chargaff, E., *Biochim. et Biophys. Acta*, **9**, 402 (1952).
[4] Wyatt, G. R., *J. Gen. Physiol.*, **36**, 201 (1952).
[5] Astbury, W. T., Symp. Soc. Exp. Biol. 1, Nucleic Acid, 66 (Camb. Univ. Press, 1947).
[6] Wilkins, M. H. F., and Randall, J. T., *Biochim. et Biophys. Acta*, **10**, 192 (1953).

Molecular Structure of Deoxypentose Nucleic Acids

While the biological properties of deoxypentose nucleic acid suggest a molecular structure containing great complexity, X-ray diffraction studies described here (cf. Astbury[1]) show the basic molecular configuration has great simplicity. The purpose of this communication is to describe, in a preliminary way, some of the experimental evidence for the polynucleotide chain configuration being helical, and existing in this form when in the natural state. A fuller account of the work will be published shortly.

The structure of deoxypentose nucleic acid is the same in all species (although the nitrogen base ratios alter considerably) in nucleoprotein, extracted or in cells, and in purified nucleate. The same linear group of polynucleotide chains may pack together parallel in different ways to give crystalline[1–3], semi-crystalline or paracrystalline material. In all cases the X-ray diffraction photograph consists of two regions, one determined largely by the regular spacing of nucleotides along the chain, and the other by the longer spacings of the chain configuration. The sequence of different nitrogen bases along the chain is not made visible.

Oriented paracrystalline deoxypentose nucleic acid ('structure *B*' in the following communication by Franklin and Gosling) gives a fibre diagram as shown in Fig. 1 (cf. ref. 4). Astbury suggested that the strong 3·4-A. reflexion corresponded to the internucleotide repeat along the fibre axis. The ~ 34 A. layer lines, however, are not due to a repeat of a polynucleotide composition, but to the chain configuration repeat, which causes strong diffraction as the nucleotide chains have higher density than the interstitial water. The absence of reflexions on or near the meridian immediately suggests a helical structure with axis parallel to fibre length.

Diffraction by Helices

It may be shown[5] (also Stokes, unpublished) that the intensity distribution in the diffraction pattern of a series of points equally spaced along a helix is given by the squares of Bessel functions. A uniform continuous helix gives a series of layer lines of spacing corresponding to the helix pitch, the intensity distribution along the nth layer line being proportional to the square of J_n, the nth order Bessel function. A straight line may be drawn approximately through

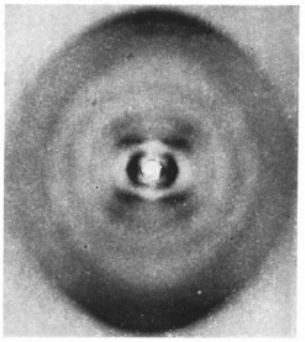


Fig. 1. Fibre diagram of deoxypentose nucleic acid from *B. coli*. Fibre axis vertical

the innermost maxima of each Bessel function and the origin. The angle this line makes with the equator is roughly equal to the angle between an element of the helix and the helix axis. If a unit repeats n times along the helix there will be a meridional reflexion (J_n^2) on the nth layer line. The helical configuration produces side-bands on this fundamental frequency, the effect[5] being to reproduce the intensity distribution about the origin around the new origin, on the nth layer line, corresponding to *C* in Fig. 2.

We will now briefly analyse in physical terms some of the effects of the shape and size of the repeat unit or nucleotide on the diffraction pattern. First, if the nucleotide consists of a unit having circular symmetry about an axis parallel to the helix axis, the whole diffraction pattern is modified by the form factor of the nucleotide. Second, if the nucleotide consists of a series of points on a radius at right-angles to the helix axis, the phases of radiation scattered by the helices of different diameter passing through each point are the same. Summation of the corresponding Bessel functions gives reinforcement for the inner-

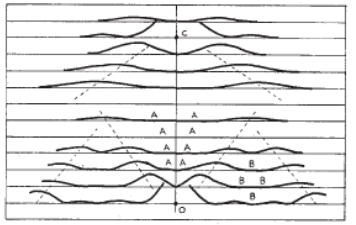


Fig. 2. Diffraction pattern of system of helices corresponding to structure of deoxypentose nucleic acid. The squares of Bessel functions are plotted about 0 on the equator and on the first, second, third and fifth layer lines for half of the nucleotide mass at 20 A. diameter and remainder distributed along a radius, the mass at a given radius being proportional to the radius. About *C* on the tenth layer line similar functions are plotted for an outer diameter of 12 A.

Proteomics 2013, *13*, 37–47 DOI 10.1002/pmic.201200233

RESEARCH ARTICLE

Quantitative proteomic analysis to decipher the differential apoptotic response of bortezomib-treated APL cells before and after retinoic acid differentiation reveals involvement of protein toxicity mechanisms

Sandrine Uttenweiler-Joseph[1,2], David Bouyssié[1,2], David Calligaris[1,2], Pierre G. Lutz[1,2], Bernard Monsarrat[1,2] and Odile Burlet-Schiltz[1,2]**

[1] CNRS, IPBS (Institut de Pharmacologie et de Biologie Structurale), Toulouse, France
[2] Université de Toulouse, UPS, IPBS, Toulouse, France

Received: June 11, 2012
Revised: August 30, 2012
Accepted: October 2, 2012

The ubiquitin-proteasome system allows the targeted degradation of proteins and plays a critical role in the regulation of many cellular processes. Proteasome inhibition is a recent antitumor therapeutic strategy and bortezomib was the first proteasome inhibitor approved for clinical use. In this study, we used the NB4 cell line to investigate the effects of bortezomib toward acute promyelocytic leukemia cells before and after retinoic acid-induced differentiation. We showed that apoptosis level after bortezomib treatment is higher in NB4 cells than in differentiated NB4 cells. To compare early protein variations upon bortezomib treatment in both NB4 cell populations, we performed a quantitative proteomic analysis based on iTRAQ peptide labeling followed by data analysis with in-house developed scripts. This strategy revealed the regulation of 14 proteins principally involved in protein stress response and apoptosis in NB4 cells after proteasome inhibition. Altogether, our results suggest that the differential level of apoptosis induced by bortezomib treatment in both NB4 cell populations could result from distinct protein toxicity level.

Keywords:
Cell biology / iTRAQ / Kinetic of protein variation / Proteasome inhibition

Additional supporting information may be found in the online version of this article at the publisher's web-site

1 Introduction

The ubiquitin-proteasome system (UPS) allows targeted protein degradation and represents an essential cellular process that contributes to the regulation of many cellular mechanisms including cell cycle progression, signal transduction, stress response, apoptosis and protein quality control. A dysfunction of this system can lead to several pathologies like inflammation disorders, neurological diseases, and cancers.

Correspondence: Dr. Sandrine Uttenweiler-Joseph, IPBS, 205 route de Narbonne, BP64182, F-31077 Toulouse Cedex, France
E-mail: Sandrine.Uttenweiler@ipbs.fr
Fax: +33 5 61 17 59 94

Abbreviations: AML, acute myeloid leukemia; **APL**, acute promyelocytic leukemia; **BZ**, bortezomib; **FDR**, false discovery rate; **MM**, multiple myeloma; **mRNP**, messenger ribonucleoprotein; **NB4 + RA**, RA-differentiated NB4 cells; **PARP**, poly(ADP-ribose) polymerase; **RA**, retinoic acid; **SCX**, strong cation exchange; **SG**, stress granules; **SR**, stress response; **UPR**, unfolded protein response; **UPS**, ubiquitin-proteasome system

*Current address: Dr. David Calligaris, Mass Spectrometry Laboratory, GIGA-Research, University of Liege, 4000 Liege, Belgium.
**Additional corresponding author: Dr. Odile Burlet-Schiltz, E-mail: Odile.Schiltz@ipbs.fr

The UPS also represents a pharmacological target in cancer therapy as illustrated by the use of proteasome inhibitors to treat several malignancies [1]. Bortezomib (Velcade®, formerly PS-341) was the first proteasome inhibitor approved by the FDA for the treatment of multiple myeloma (MM) and relapsed mantle cell lymphoma [2, 3]. This agent and more recent proteasome inhibitors like carfilzomib or NPI-0052 are currently being evaluated in combination with other agents in patients with solid tumors [4, 5].

Proteasome inhibitors have antiproliferative or proapoptotic activity against cancer cells through multiple mechanisms like stabilization of cell-cycle regulators and proapoptotic factors, induction of ER stress, activation of the unfolded protein response (UPR), inhibition of the nuclear factor kappa B inflammatory pathway, and increased generation of reactive oxygen species [6]. The in vitro and in vivo effects of proteasome inhibitors are tumor-dependent explaining the different therapeutic efficacy of these drugs. Many studies where performed to decipher these mechanisms but most of them where targeted on specific pathways and only few global transcriptomic and/or proteomic analyses were performed to discover new and unexpected modes of action [7–9] or to explain the sensitivity variation to proteasome inhibition of different cell lines from a same cancer [10].

Another important feature of proteasome inhibitors is their selectivity for tumor cells in patients. For example, MM cell lines are up to 40 times more sensitive to the proapoptotic effects of bortezomib (BZ) than are peripheral blood mononuclear cells from healthy individuals [11]. The exact mechanisms explaining this selectivity are not yet known but several hypotheses were already proposed based on the general idea that proteasome inhibitor sensitivity is linked to proliferation and/or deregulated cell cycle progression [6, 12].

In the present study, we aimed to give new insights into the differential effects of BZ toward malignant and mature hematopoietic cells from the same lineage. The NB4 cell line was used as a model because these acute promyelocytic leukemia (APL) cells are able to differentiate along the granulocytic pathway when exposed to retinoic acid (RA) leading to phenotypically mature neutrophil granulocytes [13, 14]. The level of apoptosis measured after BZ treatment proved that NB4 cells are more sensitive to proteasome inhibition than the RA-differentiated NB4 (NB4 + RA) cells. A quantitative proteomic analysis using iTRAQ technology was set up to compare the protein variations upon BZ treatment in NB4 and NB4 + RA cells. To decipher the early events leading to the apoptotic cascade initiation, the protein variations were followed at several time points thanks to the multiplex iTRAQ reagents that are ideally suited for time course studies. Based on our past experience with the development of the MFPaQ software to extract quantitative data from isotopic labeling experiments using either ICAT or SILAC methods [15], in-house scripts were elaborated to analyze the iTRAQ quantitative data. The results obtained suggest that NB4 cells are more affected by protein toxicity than NB4 + RA cells after BZ treatment.

2 Materials and methods

2.1 Cell culture and differentiation

Promyelocytic NB4 cells [13] were cultured in RPMI 1640 medium (Invitrogen) with 10% fetal bovine serum (PAA Laboratories), 2 mM glutamine and 1% penicillin-streptomycin (Invitrogen). Cells were grown in a humidified atmosphere at 37°C and 5% CO_2. Cell viability was assessed by standard trypan blue dye exclusion assay. Exponentially growing NB4 cells were seeded at 2×10^5 cells/mL 16 h prior all-*trans* RA treatment (Sigma-Aldrich, St Louis, MO; final concentration of 10^{-6} M). Differentiation was assessed by the percentage of nitro blue tetrazolium (Sigma)-positive cells.

2.2 Apoptosis analysis by flow cytometry

NB4 cells in logarithmic-phase growth or after different days of RA-induced differentiation were treated or not (control) with 0.1 nM to 1 μM BZ for up to 48 h. BZ (Velcade®) was generously provided by Millennium Pharmaceuticals Inc. (Cambridge, MA, USA). Apoptotic cells were assessed using the Annexin V-FITC/propidium iodide detection kit supplied by BD PharmingenTM (San Jose, CA, USA) as described previously [16].

2.3 Cell lysates, in-solution digestion, and iTRAQ labeling

Duplicate samples (a and b) of NB4 cells in logarithmic-phase growth or after 3 days of RA-induced differentiation (NB4 + RA) were treated or not (control) with 10 nM BZ. Cells were harvested by centrifugation before (0 h) and 6, 12, and 24 h after treatment. Cells were washed three times with ice cold PBS, frozen in liquid nitrogen, and stored at −80°C until further use. The harvested cells were resuspended in lysis buffer (10 mM HEPES, pH 7.5, 10 mM KCl, 1 mM $MgCl_2$) containing protease inhibitors (Roche, Indianapolis, Indiana, USA). Crude cell extracts were centrifuged for 10 min at 800 × g and the resulting supernatants were centrifuged at 100 000 × g for 1 h. The latter supernatants correspond to the cytosolic extracts and protein concentration was determined using the BioRad Protein Assay (BioRad, Hercules, CA, USA). A total of 100 μg of protein per each time point was used for iTRAQ labeling. Triethylammonium bicarbonate and *n*-dodecyl β-D-glucopyranoside were added to each sample to reach a final concentration of 0.5 M and 0.01%, respectively. Proteins were then reduced and alkylated according to the iTRAQ kit manufacturer's instruction (Applied Biosystems). Samples were digested with trypsin (Sequence grade Modified, Promega, Madison, WI, USA) using 1:50 ratio at 37°C overnight. Labeling with the iTRAQ reagents was performed according to manufacturer's instructions and as detailed in Supporting Information Table 1. After isobaric

Where is the knowledge we have lost in information?

T. S. ELIOT, CHORUSES FROM *THE ROCK* (1934)

Why is TDM important?

- Publishers: enrich published content, add value for customers, develop new products
- Researchers: identify new hypotheses, discover new patterns, facts, knowledge
- Corporate R&D (e.g. pharma): as above, accelerate drug discovery and development, maximise value of information spend
- Commercial users: develop new products based on cross-publisher mining & entity extraction – patents, drug discovery, product information

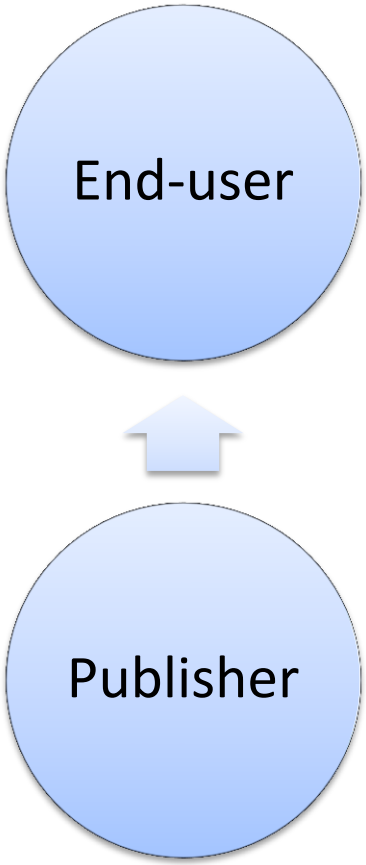
Ahead of us, there is merely … more stuff

WILLIAM GIBSON, *DISTRUST THAT PARTICULAR FLAVOR*

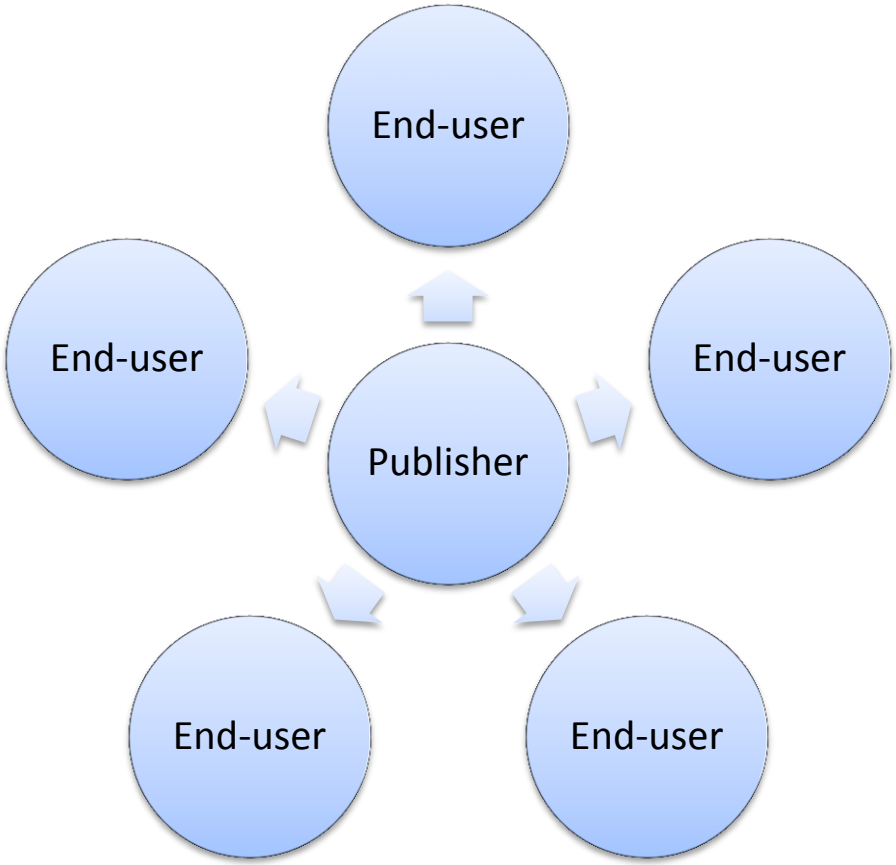
Barriers

- Access: how can users get hold of content for text mining?
- Content formats: no standard cross-publisher format
- Evaluation: understanding user needs & use-cases
- Uncertainty: what is allowed by law? What use can be made of TDM output?
- Business models: lack of established business/pricing models, e.g. access to unsubscribed content
- Scale: defining & managing demand – bilateral licensing unlikely to be scalable

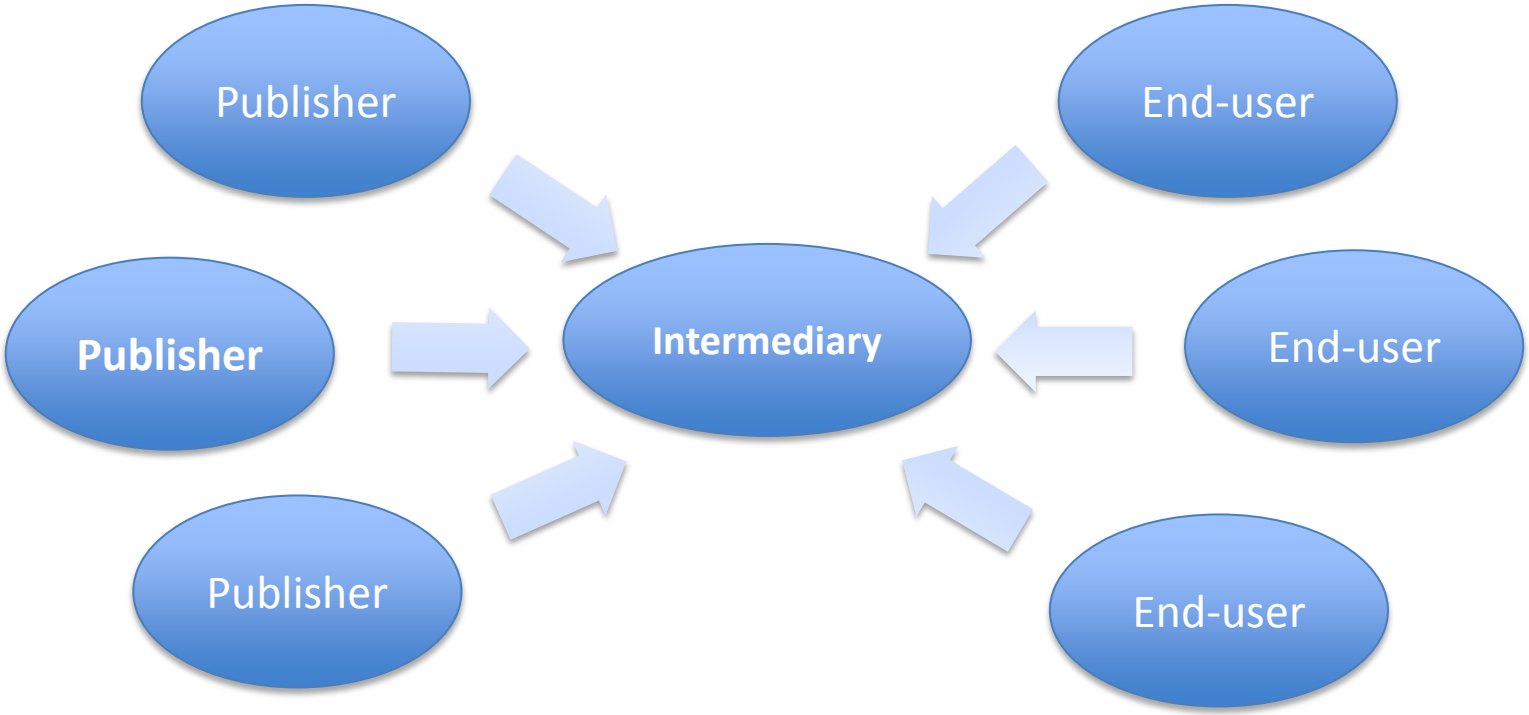
Intermediary services



Intermediary services



Intermediary services



Industry initiatives

- STM-PDR: model license for pharma
- PLS: clearing-house for requests
- Publishers Association: click-through license
- Copyright Clearance Center: pilot TDM platform
- CrossRef: ‘Prospect’ TDM solution

Intermediary services - benefits

- Single point of access and delivery
- Standard licensing terms for all end-users
- Speed and ease of access
- Extensible and scalable
- Potential to cover long tail of publishers & end-users
- Enables confidential access (essential for pharma)

Potential outcomes

- Gain a better understanding of user & market needs
- Develop services that manage content access and delivery in a scalable and extensible manner
- Eliminate needs for much bilateral licensing & individual negotiation
- Potential to develop new business models for content access (e.g. unsubscribed content)
- Solutions allow for extension into the academic research/library space – solves further significant access issues

WILEY

Duncan Campbell
dcampbell@wiley.com