La TEI dans ...

ISTEX

Stéphanie Gregorio

Inist-CNRS

stephanie.gregorio@inist.fr

TEI-ISTEX : Ce que nous allons aborder

- ❏ Présentation d'**ISTEX**
- ❏ Schéma **ODD-ISTEX**
- ❏ **Transformations** en TEI
- ❏ **Annotations** en TEI

ISTEX
L'excellence documentaire pour tous

« Construire le socle de la bibliothèque scientifique numérique nationale. »

ISTEX …

Initiative d'excellence en Information Scientifique et Technique

- **Accès en ligne aux collections rétrospectives** de la littérature scientifique dans toutes les disciplines
- Pour les abonnés de la communauté de l'**enseignement supérieur et de la recherche**

https://www.istex.fr/

ISTEX aujourd'hui, c'est :

26
corpus éditeurs

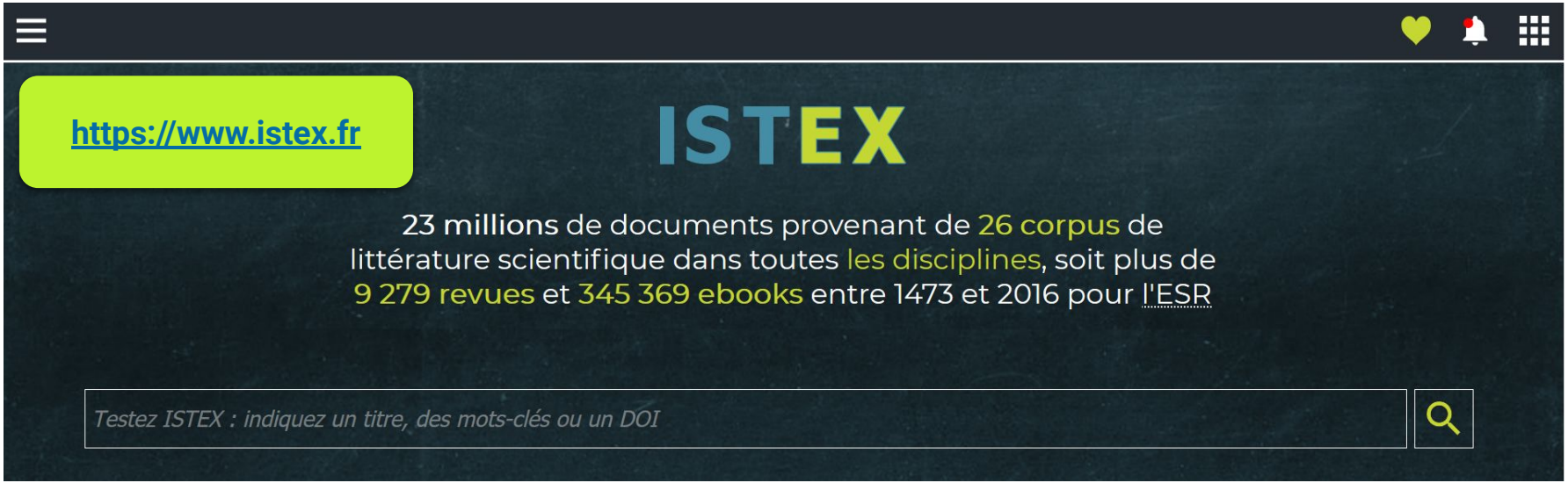
23 027 718
documents

9 279
revues

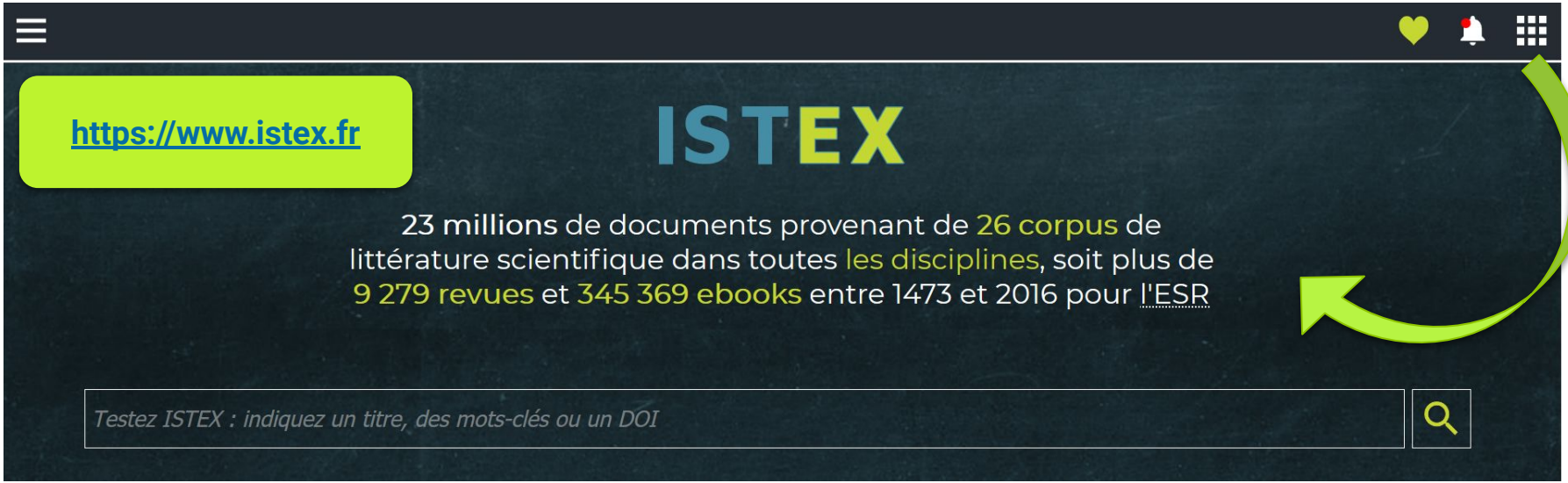
345 369
monographies

De 1473 à 2016

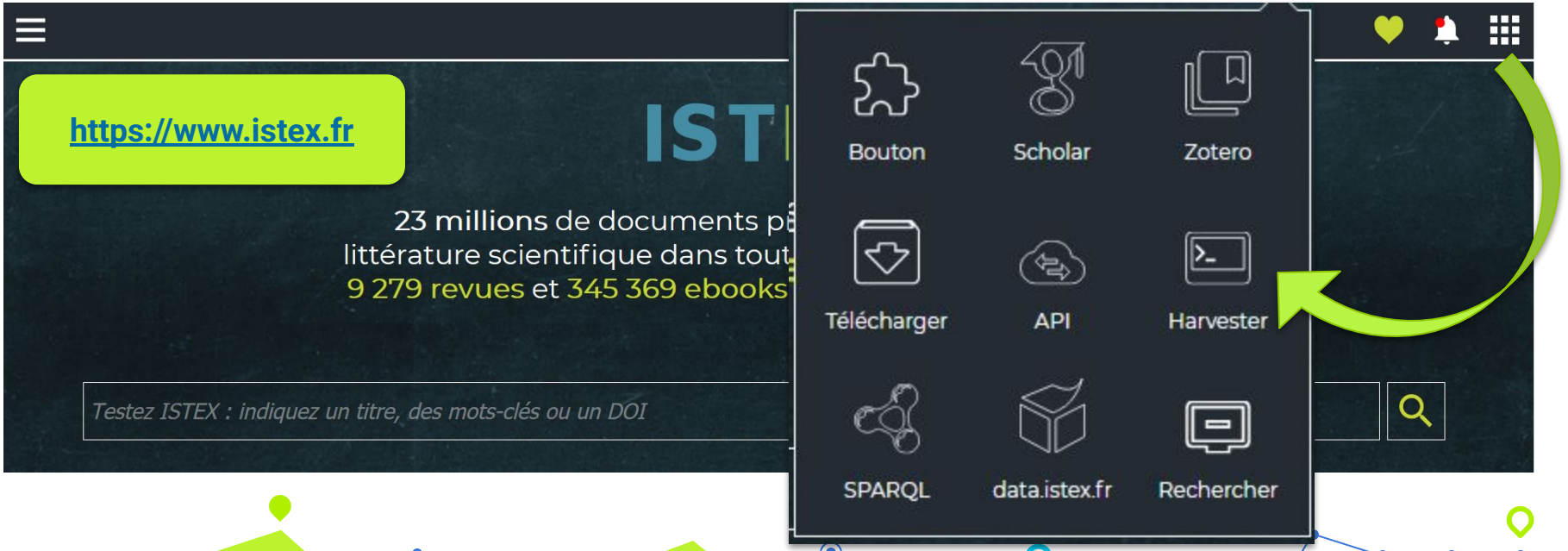
ISTEX : quels services ?



ISTEX : quels services ?



ISTEX : quels services ?



ISTEX : quels services ?

Où interroger ISTEX ?

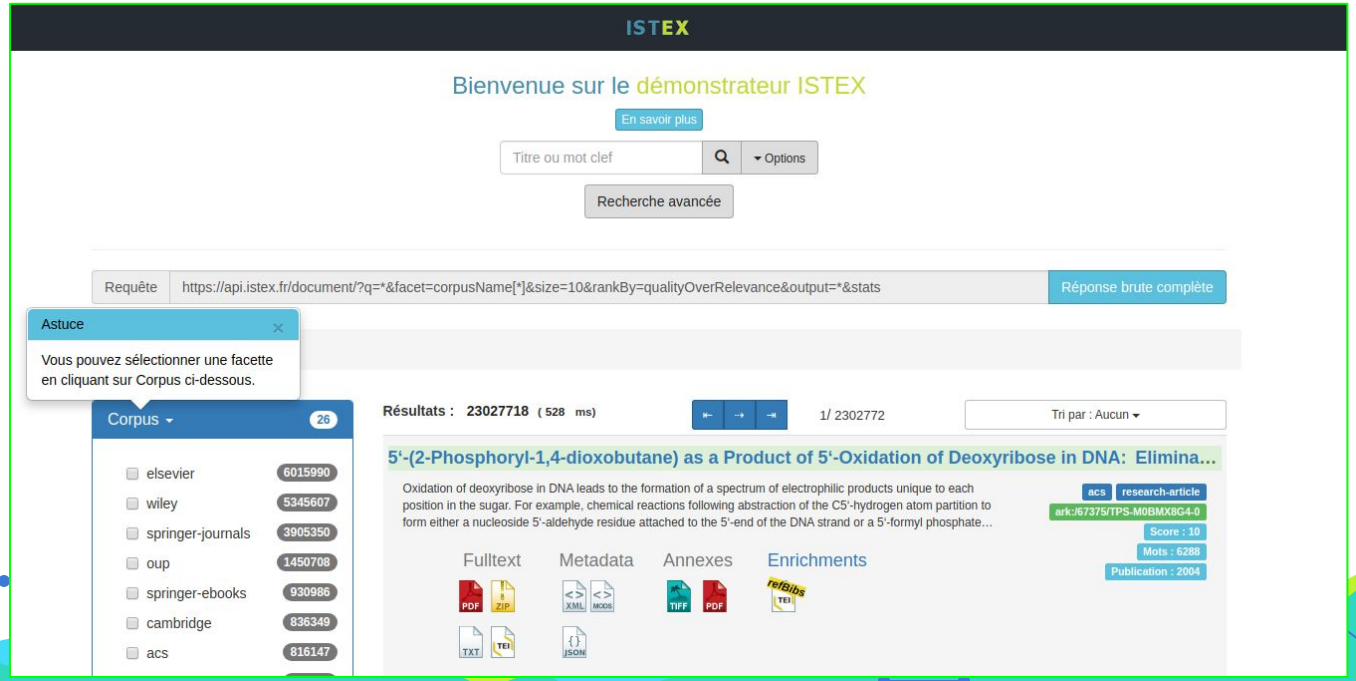
ISTEX
L'excellence documentaire pour tous

"Construire le socle de la bibliothèque scientifique numérique nationale."

Le démonstrateur est une interface à **vocation pédagogique** branchée sur l'API ISTEX qui permet de :

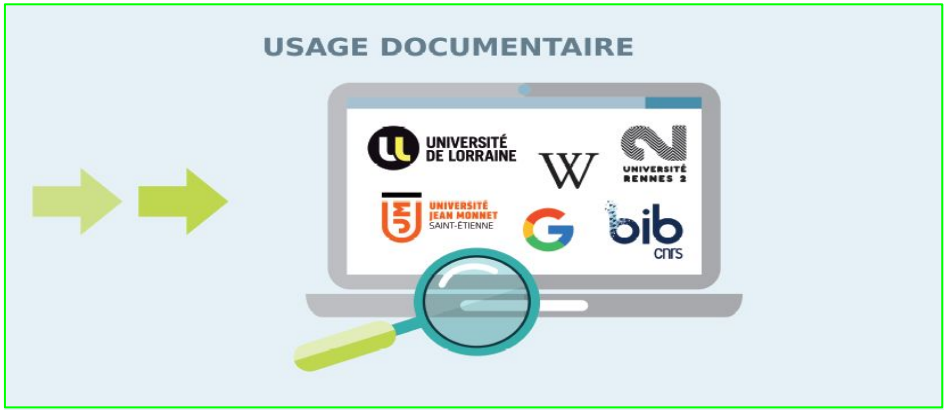
- Construire sa requête (en mode simple ou en mode avancé)
- Visualiser et filtrer les résultats

https://demo.istex.fr



ISTEX : quels usages ?

L’un classique,



… l’autre plus avancé !

La TEI dans ISTEX

Schéma TEI - ISTEX

https://tei-istex.fr

1

Schéma TEI-ISTEX = ODD-ISTEX
> modélisation des données

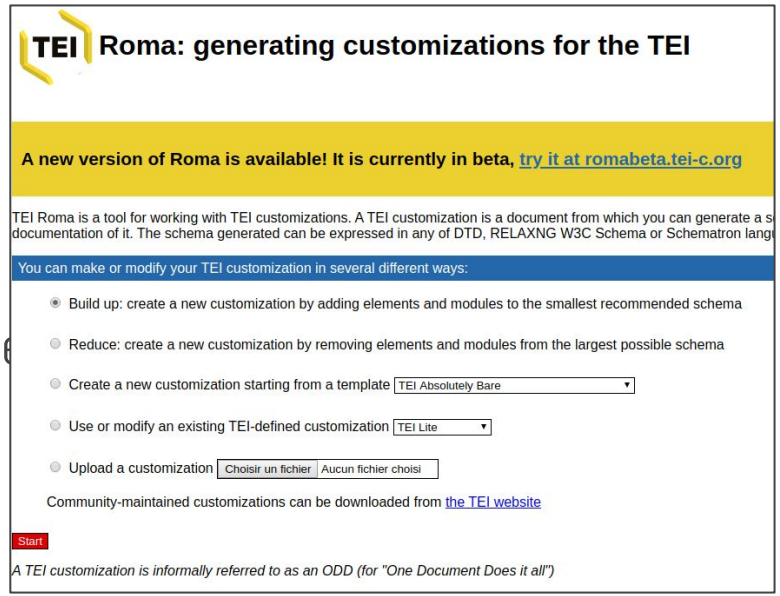
- ODD (**O**ne **D**ocument **D**oes it all) ⟶ pour personnaliser la TEI
- Implémentations techniques du modèle
    - De façon générale: correspondre aux Guidelines TEI
    - De façon particulière: répondre à des besoins spécifiques
- Adapter les modules de la TEI en fonction des cas rencontrés

Pour en savoir plus sur ce qu'est un ODD ⟶ https://halshs.archives-ouvertes.fr/cel-01706530/file/06_TEI_ODD_Camps_20170202.pdf

Schéma TEI-ISTEX = ODD-ISTEX > modélisation des données

- Génération grâce à l'outil ROMA
  - Création d'un schéma TEI **adapté**
  - Prise en charge de la **structuration** de abstracts
  - **Exports** de schémas XSD, relaxNG

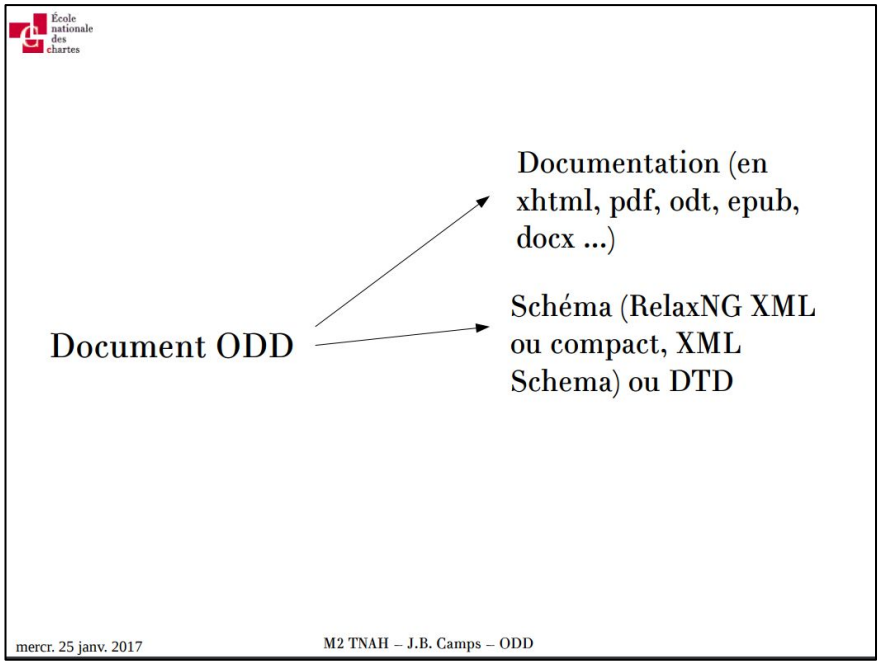
...



https://roma2.tei-c.org/

Schéma TEI-ISTEX = ODD-ISTEX > modélisation des données

- Génération grâce à l'outil ROMA
  - Export de la **documentation**



Source: https://halshs.archives-ouvertes.fr/cel-01706530/file/06_TEI_ODD_Camps_20170202.pdf

Schéma TEI pour le projet ISTEX

Mise à disposition

Schéma ODD-ISTEX	-
Disponible sur GitHub	-
Autres formats disponibles à télécharger	-

Schéma ODD-ISTEX
Codes
Schéma XSD …

Table of contents

- Guidelines personnalisés (version Bêta) pour le format TEI-ISTEX

extrait de l'outil ROMA

https://tei.istex.fr

pub2TEI-ISTEX

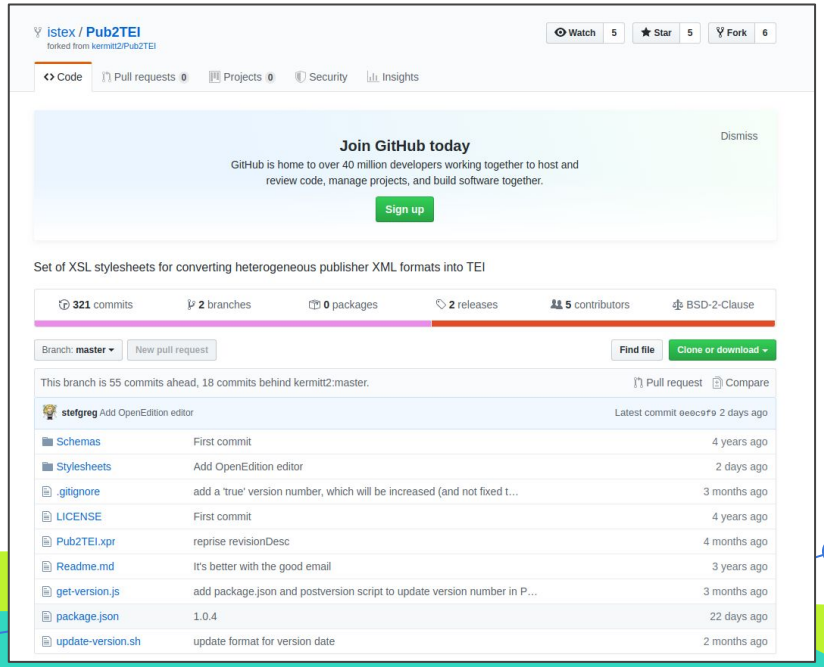
https://github.com/istex/Pub2TEI

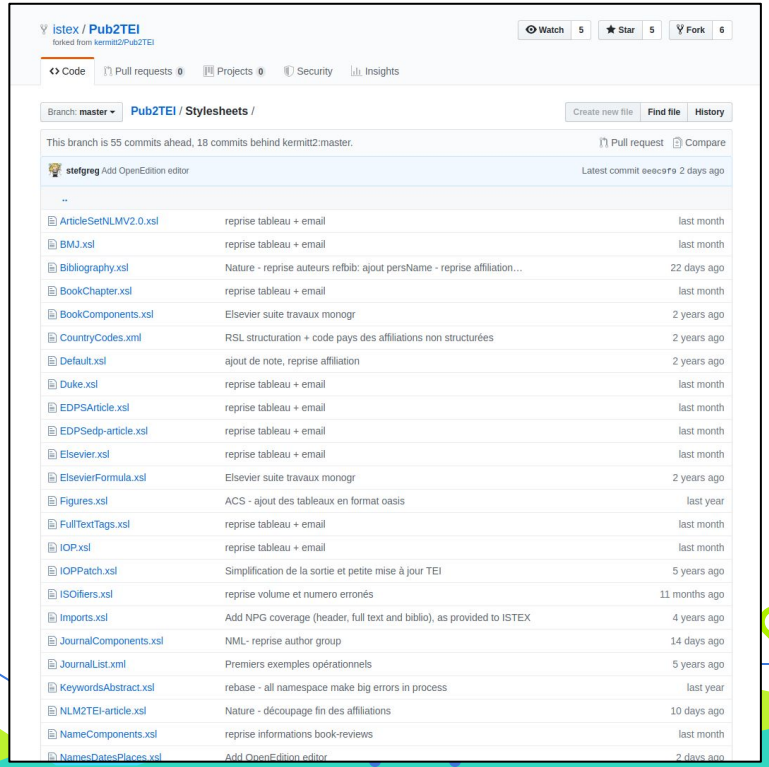
2

pub2TEI-ISTEX > modélisation des données

- Développé au départ dans le cadre du projet européen PEER par L. Romary et P. Lopez
- Propose un **ensemble de feuilles de style** → conversion documents XML en **format TEI**
- Déploiement et adaptation pour **ISTEX** *(en cours)*
  - Entrainements, développements

https://github.com/istex/pub2tei





pub2TEI-ISTEX < modélisation>

PDF

articles

A novel protein–mineral interface

Dmitriy Alexeev[1], Haizhong Zhu[2], Maolin Guo[2,3], Weiqing Zhong[2,4], Dominic J.B. Hunter[2], Weiping Yang[2,4], Dominic J. Campopiano[2] and Peter J. Sadler[2]

Published online 24 February 2003; corrected 10 March 2003 (details online); doi:10.1038/nsb903

Transferrins transport Fe^{3+} and other metal ions in mononuclear-binding sites. We present the first evidence that a member of the transferrin superfamily is able to recognize multi-nuclear oxo-metal clusters, small mineral fragments that are the most abundant forms of many metals in the environment. We show that the ferric ion–binding protein from *Neisseria gonorrhoeae* (nFbp) readily binds clusters of Fe^{3+}, Ti^{4+}, Zr^{4+} or Hf^{4+} in solution. The 1.7 Å resolution crystal structure of Hf–nFbp reveals three distinct types of clusters in an open, positively charged cleft between two hinged protein domains. A di-tyrosyl cluster nucleation motif (Tyr195-Tyr196) is situated at the bottom of this cleft and binds either a trinuclear oxo-Hf cluster, which is capped by phosphate, or a pentanuclear cluster, which in turn can be capped with phosphate. This first high-resolution structure of a protein–mineral interface suggests a novel metal-uptake mechanism and provides a model for protein-mediated mineralization/dissimilation, which plays a critical role in geochemical processes.

Organisms have developed sophisticated mechanisms for the uptake of Fe^{3+} and other metal ions from the environment. Metals in mineral form are used as sources of energy by geobacteria[1], which are critical in the control of mineral growth and phase transformations[2]. The recognition of minerals by microbial proteins is of current interest because microbes can affect the chemistry and distribution of nearly all elements in the periodic table[3,4]. Little structural information is currently available about the interactions between proteins and metal clusters during the initial stages of protein-mediated mineralization and mineral acquisition by bacteria. Here we report the surprising discovery that an iron-binding protein in the transferrin superfamily can recognize a variety of metal clusters in its iron-binding cleft, a finding that has potential implications for our understanding of metal uptake mechanisms.

Many pathogenic microorganisms require iron for virulence[5]. However, iron acquisition is a major problem for them in aerobic environments because aquated Fe^{3+} ions are highly acidic and readily form hydroxide- and oxide-bridged polymers and insoluble FeO(OH) (rust) at neutral or alkaline pH[6]. Therefore, pathogenic Gram-negative bacteria, such as *Neisseria gonorrhoeae*, *Neisseria meningitidis* and *Haemophilus influenzae*, synthesize the iron-binding protein Fbp to capture iron presented at the outer membrane and transport it across the periplasm. Fbp, a single polypeptide chain of 309 amino acids, shows the same polypeptide topology as the two lobes of serum transferrin and lactoferrin despite a lack of significant sequence similarity[7,8]. The Fe^{3+}-binding site is in an interdomain cleft — a 'Venus fly-trap' — that is open in the apo protein and closed when Fe^{3+} is bound. Fbp has been reported to bind one Fe^{3+} strongly but reversibly[9], as does each lobe of transferrin[10]. We sought to determine whether other metals bind strongly to Fbp from *N. gonorrhoeae* (nFbp) as a possible basis for the design of novel metalloantibiotics.

Multinuclear metal binding

The group four metal ions Ti^{4+}, Zr^{4+} and Hf^{4+}, similar to Fe^{3+}, are highly acidic and, therefore, expected to bind strongly to transferrin[11]. They have a high affinity for oxygen ligands — for example, hydroxide, oxide, phosphate and phenolate — and can be found along with Fe^{3+} in the same geological minerals[12]. Ti^{4+} has antibacterial properties[13] and binds strongly to human transferrin[14]. There is interest in the binding of Hf^{4+} to transferrin because of its chemical similarity to Pu^{4+}; injected Hf^{4+} is almost totally bound to transferrin in animal serum[15].

The molar ratio of protein:Fe:phosphate (P) in isolated recombinant holo-nFbp overexpressed in the *Escherichia coli* periplasm is 1:1:1 (Table 1). Addition of an equimolar amount of Hf^{4+} (as the mononuclear complex $[Hf(NTA)_2]^{2-}$) to holo-nFbp led to displacement of 47% of the bound Fe (monitored by the intensity decrease of the tyrosinate-to-Fe^{3+} charge-transfer band at 481 nm in 100 mM NaCl, 4 mM phosphate and 25 mM $NaHCO_3$, pH 7.4, at 310 K). To fully load nFbp with Hf, apo nFbp was reacted with a 50-fold molar excess of the mononuclear chelated Hf^{4+} complex. Unexpectedly, the Hf–nFbp product contained 4.0 Hf and 1.4 P per mol protein (Table 1). The presence of bound phosphate was confirmed by ^{31}P-NMR spectroscopy (new peaks shifted downfield from free phosphate by 2–3 p.p.m.; data not shown). Other metal ions, including Fe^{3+} itself, yielded similar binding patterns. Treatment of apo nFbp with excess chelated Fe^{3+} in the presence of phosphate gave rise to 2–3 mol Fe per mol protein (Table 1). Attempts to remove the excess bound Fe^{3+} by treatment with phosphate ($FePO_4$ is insoluble) did not restore the 1:1 Fe:P ratio (Table 1). Treatment of native mono-ferric holo-nFbp with a large excess of chelated Fe^{3+} did not lead to additional binding of Fe, although the bound phosphate was readily removed. The pathway for multinuclear loading, therefore, does not seem to involve the mononuclear form of the protein[2], which is isolated from bacteria.

Conformational heterogeneity in crystals

We crystallized and solved the three-dimensional structure of Hf–nFbp, containing 4.8 Hf^{4+} per mol protein, at 1.7 Å resolution by molecular replacement (Table 2). Hemihedral twinning

[1]Institute of Cell and Molecular Biology, Michael Swann Building, University of Edinburgh, Mayfield Road, Edinburgh EH9 3JR, U.K. [2]School of Chemistry, University of Edinburgh, West Mains Road, Edinburgh EH9 3JJ, UK. [3]Current address: Department of Molecular Biology and Biochemistry, University of California, Irvine, California 92697, USA. [4]Current address: School of Pharmacy, Second Military Medicine University, Shanghai 200433, China.

Correspondence should be addressed to P.J.S. e-mail: P.J.Sadler@ed.ac.uk

Format XML éditeur

```
<?xml version="1.0"?>
<article id="nsb903" language="eng" publish="issue" relation="no">
    <pubfm>
        <jtl>Nature Structural Biology</jtl>
        <vol>10</vol>
        <iss>4</iss>
        <idt>200304</idt>
        <categ id="af"/>
        <pp>
            <spn>297</spn>
            <epn>302</epn>
            <cnt>6</cnt>
        </pp>
        <issn type="print">1072-8368</issn>
        <cpg>
            <cpy>2003</cpy>
            <cpn>Nature Publishing Group</cpn>
        </cpg>
        <nvid extrefid="nsb0403-240"/>
        <chghst>
            <chg chgtype="erratum">
                <chgperson>
                    <fnm>Susan</fnm>
                    <snm>Kline</snm>
                </chgperson>
                <chgdate year="2003" month="03" day="07"/>
                <chgloc>Author affiliations</chgloc>
                <chgmade>This was incorrect in AOP version but corrected in print. Revised author
                    affiliation numbers and affiliation information as per the note.</chgmade>
                <chgreason>Editorial error.</chgreason>
            </chg>
        </chghst>
        <doi>10.1038/nsb903</doi>
    </pubfm>
    <fm>
        <atl>A novel proteinn-mineral interface</atl>
        <aug>
            <au>
                <fnm>Dmitriy</fnm>
                <snm>Alexeev</snm>
                <inits>D</inits>
                <orf rid="a1"/>
            </au>
            <au> [5 lines]
            <au> [6 lines]
            <au> [6 lines]
            <au> [5 lines]
            <au> [6 lines]
            <au> [5 lines]
            <cau> [6 lines]
        <aff><oid id="a1" /><org>Institute of Cell and Molecular Biology, Michael Swann Building,
                University of Edinburgh</org>, <street>Mayfield Road</street>,
                <cty>Edinburgh</cty>
            <zip>EH9 3JR</zip>, <cny>UK</cny>.</aff>
        <aff><oid id="a2"/><org>School of Chemistry, University of Edinburgh</org>, <street>West
                Mains Road</street>, <cty>Edinburgh</cty>
            <zip>EH9 3JJ</zip>, <cny>UK</cny>.</aff>
        <aff><oid id="a3"/>Current address: <org>Department of Molecular Biology and
                Biochemistry, University of California</org>, <cty>Irvine</cty>,
                <st>California</st>
            <zip>92697</zip>, <cny>USA</cny>.</aff>
        <aff><oid id="a4"/>Current address: <org>School of Pharmacy, Second Military Medicine
                University</org>, <cty>Shanghai</cty>
            <zip>200433</zip>, <cny>China</cny>.</aff>
```

Article Journal *Nature*
https://api.istex.fr/ark:/67375/GT4-LHM354SF-Z/fulltext.pdf?sid=istex-api-demo

pub2TEI-ISTEX

< modélisation des affiliations semi-structurées>

XML Editor

nature publishing group npg

Article Journal *Nature*
https://api.istex.fr/ark:/67375/GT4-LHM354SF-Z/fulltext.pdf?sid=istex-api-demo

Format XML éditeur

```
<fm>
    <atl>A novel proteinn-mineral interface</atl>
    <aug>
        <au>
            <fnm>Dmitriy</fnm>
            <snm>Alexeev</snm>
            <inits>D</inits>
            <orf rid="a1"/>
        </au>
        <aff>
            <oid id="a1"/>
            <org>Institute of Cell and Molecular Biology, Michael Swann Building,
                University of Edinburgh</org>,
            <street>Mayfield Road</street>,
            <cty>Edinburgh</cty>
            <zip>EH9 3JR</zip>,
            <cny>UK</cny>.</aff>
    </aug>
    <hst>
        <re year="2002" month="09" day="04"/>
        <acc year="2003" month="01" day="06"/>
```

Reformatage en TEI

```
<biblStruct>
   <analytic>
      <title level="a" type="main" xml:lang="en">A novel proteinn-mineral interface</title>
      <author xml:id="author-0000">
         <persName>
            <forename type="first">Dmitriy</forename>
            <surname>Alexeev</surname>
            <forename full="init">D</forename>
         </persName>
         <affiliation>
            <orgName type="institution">Institute of Cell and Molecular Biology</orgName>
            <orgName>Michael Swann Building</orgName>
            <orgName type="institution">University of Edinburgh</orgName>
            <address>
               <street>Mayfield Road</street>
               <postCode>EH9 3JR</postCode>
               <settlement>Edinburgh</settlement>
               <country key="GB" xml:lang="en">UNITED KINGDOM</country>
            </address>
         </affiliation>
      </author>
```

Découpage de façon automatique des organismes afin de les baliser

ISO 3166-1 alpha-2 + verbalisation

pub2TEI-ISTEX
<restructuration des affiliations *non structurées*>

WILEY

Format XML éditeur

```
<titleGroup>
    <title type="main">Electron transfer at the microbe–mineral interface: a grand
        challenge in biogeochemistry</title>
    <title type="shortAuthors">J. K. FREDRICKSON AND J. M. ZACHARA</title>
    <title type="short">Electron transfer to minerals</title>
</titleGroup>
<creators>
    <creator creatorRole="author" xml:id="cr1" affiliationRef="#aff-1-1">
        <personName>
            <givenNames>J. K.</givenNames>
            <familyName>FREDRICKSON</familyName>
        </personName>
    </creator>
    <creator creatorRole="author" xml:id="cr2" affiliationRef="#aff-1-1">
        <personName>
            <givenNames>J. M.</givenNames>
            <familyName>ZACHARA</familyName>
        </personName>
    </creator>
</creators>
<affiliationGroup>
    <affiliation xml:id="aff-1-1" countryCode="US">
        <unparsedAffiliation>Pacific Northwest National Laboratory, PO Box 999,
            Richland, WA 99352, USA</unparsedAffiliation>
    </affiliation>
</affiliationGroup>
<abstractGroup>
    <abstract type="main" xml:lang="en">
        <title type="main">ABSTRACT</title>
        <p>The interplay between microorganisms and minerals is a complex and dynamic
            process that has sculpted the geosphere for nearly the entire history of the
            Earth. The work of Dr Terry Beveridge and colleagues provided some of the
            first insights into metal–microbe and mineral–microbe interactions and
```

Reformatage en TEI

```
31  <biblStruct type="article">
32    <analytic>
33      <title level="a" type="main">Electron transfer at the microbe–mineral interface: a grand cha
34      <title level="a" type="short">Electron transfer to minerals</title>
35      <author xml:id="author-0000">
36        <persName>
37          <forename type="first">J. K.</forename>
38          <surname>FREDRICKSON</surname>
39        </persName>
40        <affiliation>
41          <orgName type="laboratory">Pacific Northwest National Laboratory</orgName>
42          <address>
43            <addrLine>PO Box 999</addrLine>
44            <addrLine>Richland</addrLine>
45            <addrLine>WA 99352</addrLine>
46            <country key="US">UNITED STATES</country>
47          </address>
          </affiliation>
49      </author>
50      <author xml:id="author-0001">
51        <persName>
52          <forename type="first">J. M.</forename>
53          <surname>ZACHARA</surname>
54        </persName>
55        <affiliation>
56          <orgName type="laboratory">Pacific Northwest National Laboratory</orgName>
57          <address>
58            <addrLine>PO Box 999</addrLine>
59            <addrLine>Richland</addrLine>
60            <addrLine>WA 99352</addrLine>
61            <country key="US">UNITED STATES</country>
62          </address>
63        </affiliation>
```

ISO 3166-1 alpha-2 + verbalisation

Article Journal *Geobiology* https://api.istex.fr/ark:/67375/WNG-M06V00BL-F/fulltext.pdf?sid=istex-api-demo

pub2TEI-ISTEX <modélisation du corps du texte>

Article Journal *Nature*
https://api.istex.fr/ark:/67375/GT4-LHM354SF-Z/fulltext.pdf?sid=istex-api-demo

Format XML éditeur

```
<bdy>
   <p>Organisms have developed sophisticated mechanisms for the uptake of Fe<super>3+</super>
      and other metal ions from the environment. Metals in mineral form are used as sources of
      energy by geobacteria<bibr rid="b1"/>, which are critical in the control of mineral
      growth and phase transformations<bibr rid="b2"/>. The recognition of minerals by
      microbial proteins is of current interest because microbes can affect the chemistry and
      distribution of nearly all elements in the periodic table<bibr rid="b3 b4"/>. Little
      structural information is currently available about the interactions between proteins
      and metal clusters during the initial stages of protein-mediated mineralization and
      mineral acquisition by bacteria. Here we report the surprising discovery that an
      iron-binding protein in the transferrin superfamily can recognize a variety of metal
      clusters in its iron-binding cleft, a finding that has potential implications for our
      understanding of metal uptake mechanisms.</p>
   <p>Many pathogenic microorganisms require iron for virulence<bibr rid="b5"/>. However, iron
      acquisition is a major problem for them in aerobic environments because aquated
         Fe<super>3+</super> ions are highly acidic and readily form hydroxide- and
      oxide-bridged polymers and insoluble FeO(OH) (rust) at neutral or alkaline pH<bibr
         rid="b6"/>. Therefore, pathogenic Gram-negative bacteria, such as <i>Neisseria
         gonorrhoeae</i>, <i>Neisseria meningitidis</i> and <i>Haemophilus influenzae</i>,
      synthesize the iron-binding protein Fbp to capture iron presented at the outer membrane
      and transport it across the periplasm. Fbp, a single polypeptide chain of 309 amino
      acids, shows the same polypeptide topology as the two lobes of serum transferrin and
      lactoferrin despite a lack of significant sequence similarity<bibr rid="b7 b8"/>. The
         Fe<super>3+</super>-binding site is in an interdomain cleft mdash; a 'Venus
      fly-trap' mdash; that is open in the apo protein and closed when Fe<super>3+</super> is
      bound. Fbp has been reported to bind one Fe<super>3+</super> strongly but
         reversibly<bibr rid="b9"/>, as does each lobe of transferrin<bibr rid="b10"/>. We
      sought to determine whether other metals bind strongly to Fbp from <i>N. gonorrhoeae</i>
      (nFbp) as a possible basis for the design of novel metalloantibiotics.</p>
   <crosshd>Multinuclear metal binding</crosshd>
   <p>The group four metal ions Ti<super>4+</super>, Zr<super>4+</super> and
         Hf<super>4+</super>, similar to Fe<super>3+</super>, are highly acidic and,
      therefore, expected to bind strongly to transferrin<bibr rid="b11"/>. They have a high
      affinity for oxygen ligands mdash; for example, hydroxide, oxide, phosphate and
      phenolate mdash; and can be found along with Fe<super>3+</super> in the same geological
         minerals<bibr rid="b12"/>. Ti<super>4+</super> has antibacterial properties<bibr
         rid="b13"/> and binds strongly to human transferrin<bibr rid="b14"/>. There is
      interest in the binding of Hf<super>4+</super> to transferrin because of its chemical
      similarity to Pu<super>4+</super>; injected Hf<super>4+</super> is almost totally bound
      to transferrin in animal serum<bibr rid="b15"/>.</p>
   <p>The molar ratio of protein:Fe:phosphate (P) in isolated recombinant holo-nFbp
      overexpressed in the <i>Escherichia coli</i> periplasm is 1:1:1 (<tablr rid="t1">Table
         1</tablr>). Addition of an equimolar amount of Hf<super>4+</super> (as the
      mononuclear complex [Hf(NTA)<sub>2</sub>]<super>2minus;</super>) to holo-nFbp led to
      displacement of 47% of the bound Fe (monitored by the intensity decrease of the
         tyrosinate-to-Fe<super>3+</super> charge-transfer band at 481 nm in 100 mM NaCl, 4
      mM phosphate and 25 mM NaHCO<sub>3</sub>, pH 7.4, at 310 K). To fully load nFbp with Hf,
      apo nFbp was reacted with a 50-fold molar excess of the mononuclear chelated
         Hf<super>4+</super> complex. Unexpectedly, the Hfndash;nFbp product contained 4.0
      Hf and 1.4 P per mol protein (<tablr rid="t1">Table 1</tablr>). The presence of bound
      phosphate was confirmed by <super>31</super>P-NMR spectroscopy (new peaks shifted
      downfield from free phosphate by 2ndash;3 p.p.m.; data not shown). Other metal ions,
      including Fe<super>3+</super> itself, yielded similar binding patterns. Treatment of apo
      nFbp with excess chelated Fe<super>3+</super> in the presence of phosphate gave rise to
      2ndash;3 mol Fe per mol protein (<tablr rid="t1">Table 1</tablr>). Attempts to remove
      the excess bound Fe<super>3+</super> by treatment with phosphate (FePO<sub>4</sub> is
      insoluble) did not restore the 1:1 Fe:P ratio (<tablr rid="t1">Table 1</tablr>).
      Treatment of native mono-ferric holo-nFbp with a large excess of chelated
         Fe<super>3+</super> did not lead to additional binding of Fe, although the bound
      phosphate was readily removed. The pathway for multinuclear loading, therefore, does not
```

Reformatage en TEI

```
<body>
   <p>Organisms have developed sophisticated mechanisms for the uptake of Fe<hi rend="superscript">3+</hi>
      and other metal ions from the environment. Metals in mineral form are used as sources of
      energy by geobacteria<ref type="bibr" target="#b1" />, which are critical in the control of mineral
      growth and phase transformations<ref type="bibr" target="#b2" />. The recognition of minerals by
      microbial proteins is of current interest because microbes can affect the chemistry and
      distribution of nearly all elements in the periodic table<ref type="bibr" target="#b3 #b4" />. Little
      structural information is currently available about the interactions between protei
      and metal clusters during the initial stages of protein-mediated mineralization and
      mineral acquisition by bacteria. Here we report the surprising discovery that an
      iron-binding protein in the transferrin superfamily can recognize a variety of meta
      clusters in its iron-binding cleft, a finding that has potential implications for our
      understanding of metal uptake mechanisms.</p>
   <p>Many pathogenic microorganisms require iron for virulence<ref type="bibr" target="#b5" />. However, iron
      acquisition is a major problem for them in aerobic environments because aquated
         Fe<hi rend="superscript">3+</hi> ions are highly acidic and readily form hydroxide- and
      oxide-bridged polymers and insoluble FeO(OH) (rust) at neutral or alkaline pH<ref type="bibr" target="#b6" />. Therefore
      synthesize the iron-binding protein Fbp to capture iron presented at the outer membrane
      and transport it across the periplasm. Fbp, a single polypeptide chain of 309 amino
      acids, shows the same polypeptide topology as the two lobes of serum transferrin and
      lactoferrin despite a lack of significant sequence similarity<ref type="bibr" target="#b7 #b8" />. The
         Fe<hi rend="superscript">3+</hi>-binding site is in an interdomain cleft mdash; a 'Venus
      fly-trap' mdash; that is open in the apo protein and closed when Fe<hi rend="superscript">3+</hi> is
      bound. Fbp has been reported to bind one Fe<hi rend="superscript">3+</hi> strongly but
         reversibly<ref type="bibr" target="#b9" />, as does each lobe of transferrin<ref type="bibr" target="#b10" />. We
      sought to determine whether other metals bind strongly to Fbp from
      (nFbp) as a possible basis for the design of novel metalloantibiotics.</p>
   <p>
      <hi rend="bold">Multinuclear metal binding</hi>
   </p>
   <p>The group four metal ions Ti<hi rend="superscript">4+</hi>, Zr<hi rend="superscript">4+</hi> and
         Hf<hi rend="superscript">4+</hi>, similar to Fe<hi rend="superscript">3+</hi>, are highly acidic and,
      therefore, expected to bind strongly to transferrin<ref type="bibr" target="#b11" />. They have a high
      affinity for oxygen ligands mdash; for example, hydroxide, oxide, phosphate and
      phenolate mdash; and can be found along with Fe<hi rend="superscript">3+</hi> in the same geological
         minerals<ref type="bibr" target="#b12" />. Ti<hi rend="superscript">4+</hi> has antibacterial properties<ref type="b
      interest in the binding of Hf<hi rend="superscript">4+</hi> to transferrin because of its chemical
      similarity to Pu<hi rend="superscript">4+</hi>; injected Hf<hi rend="superscript">4+</hi> is almost totally bound
      to transferrin in animal serum<ref type="bibr" target="#b15" />.</p>
   <p>The molar ratio of protein:Fe:phosphate (P) in isolated recombinant holo-nFbp
      overexpressed in the  periplasm is 1:1:1 (<ref type="table" target="#t1">Table
         1</ref>). Addition of an equimolar amount of Hf<hi rend="superscript">4+</hi> (as the
      mononuclear complex [Hf(NTA)<hi rend="subscript">2</hi>]<hi rend="superscri
      displacement of 47% of the bound Fe (monitored by the intensity decrease of
         tyrosinate-to-Fe<hi rend="superscript">3+</hi> charge-transfer band at
      mM phosphate and 25 mM NaHCO<hi rend="subscript">3</hi>, pH 7.4, at 310 K). To fully load nFbp with Hf,
      apo nFbp was reacted with a 50-fold molar excess of the mononuclear chelated
         Hf<hi rend="superscript">4+</hi> complex. Unexpectedly, the Hfndash;nFbp product contained 4.0
      Hf and 1.4 P per mol protein (<ref type="table" target="#t1">Table 1</ref>). The presence of bound
      phosphate was confirmed by <hi rend="superscript">31</hi>P-NMR spectroscopy (new peaks shifted
      downfield from free phosphate by 2ndash;3 p.p.m.; data not shown). Other metal ions,
      including Fe<hi rend="superscript">3+</hi> itself, yielded similar binding patterns. Treatment of apo
      nFbp with excess chelated Fe<hi rend="superscript">3+</hi> in the presence of phosphate gave rise to
      2ndash;3 mol Fe per mol protein (<ref type="table" target="#t1">Table 1</ref>). Attempts to remove
      the excess bound Fe<hi rend="superscript">3+</hi> by treatment with phosphate (FePO<hi rend="subscript">4</hi> is
      insoluble) did not restore the 1:1 Fe:P ratio (<ref type="table" target="#t1">Table 1</ref>).
      Treatment of native mono-ferric holo-nFbp with a large excess of chelated
         Fe<hi rend="superscript">3+</hi> did not lead to additional binding of Fe, although the bound
      phosphate was readily removed. The pathway for multinuclear loading, therefore, does not
      seem to involve the mononuclear form of the protein<ref type="bibr" target="#b7" />, which is isolated
```

style

Appel Ref Bib

Appel Ref Tableau

pub2TEI-ISTEX <modélisation des tableaux>

Article Journal *Nature*
https://api.istex.fr/ark:/67375/GT4-LHM354SF-Z/fulltext.pdf?sid=istex-api-demo

Format XML éditeur

```
<table tocentry="0" verified="0" frame="topbot" colsep="0" rowsep="0" orient="port"
    pgwide="0" id="t1" hcp="0" esds="0" tabcols="5" security="u">
    <title security="u">Analysis of nFbp samples</title>
    <tgroup cols="5" colsep="0" rowsep="0" align="center" charoff="50" char=""
        show="yes" security="u">
        <tbody valign="top" security="u">
            <row security="u">
                <entry rotate="0" valign="top" security="u">Sample</entry>
                <entry rotate="0" valign="top" security="u">Treatment<fnr rid="t1-fn1"
                    ><super>1</super></fnr></entry>
                <entry spanname="s3to5" rotate="0" valign="top" align="center"
                    security="u">Mol element per mol protein</entry>
            </row>
            <row security="u">
                <entry colname="3" rotate="0" valign="top" security="u">Fe</entry>
                <entry rotate="0" valign="top" security="u">P</entry>
                <entry rotate="0" valign="top" security="u">Other</entry>
            </row>
            <row security="u">
                <entry rotate="0" valign="top" security="u">Holo-native</entry>
                <entry rotate="0" valign="top" security="u">As isolated<fnr rid="t1-fn2"
                    ><super>2</super></fnr></entry>
                <entry rotate="0" valign="top" security="u">1.04 plusmn; 0.03</entry>
                <entry rotate="0" valign="top" security="u">0.98 plusmn; 0.04</entry>
            </row>
            <row security="u">
                <entry rotate="0" valign="top" security="u">Apo/Fe</entry>
                <entry rotate="0" valign="top" security="u">50 times;
                    [Fe(NTA)<sub>2</sub>]<super>3minus;</super>; 10 mM Tris, pH
                    7.8, 12 h</entry>
                <entry rotate="0" valign="top" security="u">2.4 plusmn; 0.2</entry>
                <entry rotate="0" valign="top" security="u">0.05 plusmn; 0.08</entry>
            </row>
            <row security="u">
                <entry rotate="0" valign="top" security="u">Apo/Fe/P<sub>i</sub><fnr
                    rid="t1-fn3"><super>3</super></fnr></entry>
                <entry rotate="0" valign="top" security="u">10 times;
                    [Fe(NTA)<sub>2</sub>]<super>3minus;</super> 10 mM Tris, pH 7.8,
                    then 10times; phosphate, 1.5 h</entry>
                <entry rotate="0" valign="top" security="u">1.8 plusmn; 0.1</entry>
                <entry rotate="0" valign="top" security="u">1.1 plusmn; 0.1</entry>
            </row>
            <row security="u">
                <entry rotate="0" valign="top" security="u">Holo/Fe</entry>
                <entry rotate="0" valign="top" security="u">10 times;
                    [Fe(NTA)<sub>2</sub>]<super>3minus;</super>; 10 mM Tris, pH
                    8.0, 4 h</entry>
                <entry rotate="0" valign="top" security="u">1.10 plusmn; 0.03</entry>
                <entry rotate="0" valign="top" security="u">0.20 plusmn; 0.01</entry>
            </row>
            <row security="u">
                <entry rotate="0" valign="top" security="u">Apo/Ti<fnr rid="t1-fn4"
                    ><super>4</super></fnr></entry>
                <entry rotate="0" valign="top" security="u">100 times; [Cp<sub>2</sub>
                    TiCl<sub>2</sub>] (in 10% DMSO, 100 mM NaCl); 10 mM Tris, pH
                    8.0, 5 h</entry>
                <entry rotate="0" valign="top" security="u">ndash;</entry>
                <entry rotate="0" valign="top" security="u">ndash;</entry>
                <entry rotate="0" valign="top" security="u">Ti 3.4 plusmn; 0.1</entry>
```

Reformatage en TEI

Entête

Appel Réf bas de page

Style

```
<table xml:id="t1" cols="5">
    <head>Analysis of nFbp samples</head>
    <row>
        <cell>Sample</cell>
        <cell>Treatment<ref type="footnote" target="#t1-fn1">
            <hi rend="superscript">1</hi>
        </ref>
        </cell>
        <cell>Mol element per mol protein</cell>
    </row>
    <row>
        <cell cols="3">Fe</cell>
        <cell>P</cell>
        <cell>Other</cell>
    </row>
    <row>
        <cell>Holo-native</cell>
        <cell>As isolated<ref type="footnote" target="#t1-fn2">
            <hi rend="superscript">2</hi>
        </ref>
        </cell>
        <cell>1.04 plusmn; 0.03</cell>
        <cell>0.98 plusmn; 0.04</cell>
    </row>
    <row>
        <cell>Apo/Fe</cell>
        <cell>50 times;
            [Fe(NTA)<hi rend="subscript">2</hi>]<hi rend="superscript">3minus;</hi>; 10 mM Tris, pH
            7.8, 12 h</cell>
        <cell>2.4 plusmn; 0.2</cell>
        <cell>0.05 plusmn; 0.08</cell>
    </row>
    <row>
        <cell>Apo/Fe/P<hi rend="subscript">i</hi>
            <ref type="footnote" target="#t1-fn3">
                <hi rend="superscript">3</hi>
            </ref>
        </cell>
        <cell>10 times;
            [Fe(NTA)<hi rend="subscript">2</hi>]<hi rend="superscript">3minus;</hi> 10 mM Tris, pH 7.8,
            then 10times; phosphate, 1.5 h</cell>
        <cell>1.8 plusmn; 0.1</cell>
        <cell>1.1 plusmn; 0.1</cell>
    </row>
    <row>
        <cell>Holo/Fe</cell>
        <cell>10 times;
            [Fe(NTA)<hi rend="subscript">2</hi>]<hi rend="superscript">3minus;</hi>; 10 mM Tris, pH
            8.0, 4 h</cell>
        <cell>1.10 plusmn; 0.03</cell>
        <cell>0.20 plusmn; 0.01</cell>
    </row>
    <row>
        <cell>Apo/Ti<ref type="footnote" target="#t1-fn4">
            <hi rend="superscript">4</hi>
        </ref>
        </cell>
        <cell>100 times; [Cp<hi rend="subscript">2</hi>
            TiCl<hi rend="subscript">2</hi>] (in 10% DMSO, 100 mM NaCl); 10 mM Tris, pH
```

Enrichissements ISTEX ou “annotations”

3

ISTEX > Outils d'enrichissement

Unitex

Détection de 9 types d'entités nommées (lieux géographiques, organismes financeurs, etc.)

enrichment-process.data.istex.fr/unitex

⇨ **66,5%** des docs

Grobid

Structuration des références bibliographiques

enrichment-process.data.istex.fr/grobid

⇨ **57,9%** des docs

Teeft

Indexation des articles par extraction de termes à partir du texte intégral en anglais

enrichment-process.data.istex.fr/teeft

⇨ **62,3%** des docs

Bayésien naïf (nb)

Catégorisation de l'article par apprentissage avec Pascal/Francis

enrichment-process.data.istex.fr/nb

⇨ **36,2%** des docs

Multicat

Catégorisation de la revue par appariement (WoS / Science Metrix / Scopus)

enrichment-process.data.istex.fr/wos

enrichment-process.data.istex.fr/sciencemetrix

enrichment-process.data.istex.fr/scopus

⇨ **71,1%** des docs

Textes intégraux à valeur ajoutée téléchargeables massivement

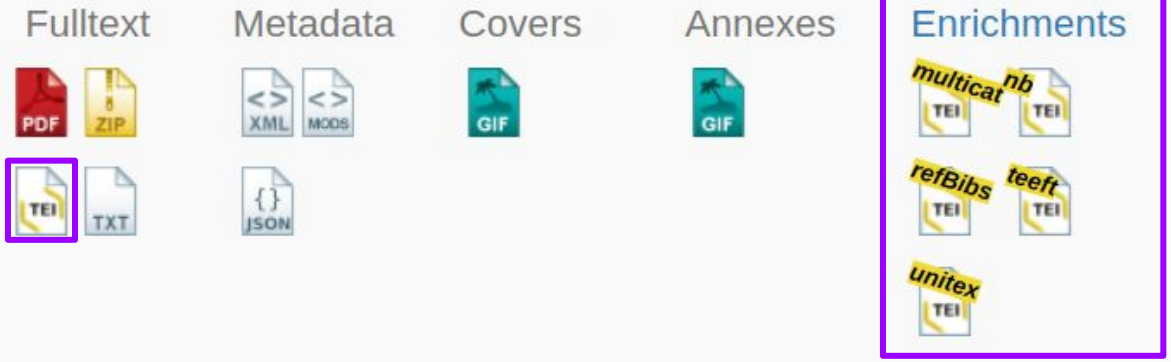
ISTEX > Résultats des Outils d'enrichissement

Article Journal *Nature*
https://api.istex.fr/ark:/67375/GT4-LHM354SF-Z/fulltext.pdf?sid=istex-api-demo

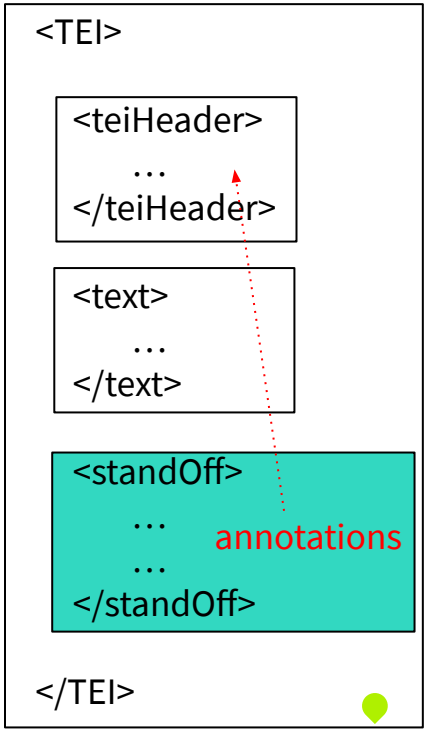
A novel protein–mineral interface

Transferrins transport Fe3+ and other metal ions in mononuclear-binding sites. We present the first evidence that a member of the transferrin superfamily is able to recognize multi-nuclear oxo-metal clusters, small mineral fragments that are the most abundant forms of many metals in the environment. We show that the ferric…

nature | research-article
ark:/67375/GT4-LHM354SF-Z
Score : 9.88
Mots : 4880
Publication : 2003



Exemple d'appariement de catégories *(niveau revue par ISSN)*




```
▼<ns1:standOff>
  ▶<teiHeader>...</teiHeader>
  ▼<ns1:listAnnotation type="rd-multicat">
    ▼<annotationBlock xmls="https//www.tei-c.org/ns/1.0">
      ▼<keywords change="#rd-multicat" resp="#istex" scheme="#science-metrix">
         <term level="1">health sciences</term>
         <term level="2">biomedical research</term>
         <term level="3">biophysics</term>
       </keywords>
      ▼<keywords change="#rd-multicat" resp="#istex" scheme="#scopus">
         <term level="1">Life Sciences</term>
         <term level="2">Biochemistry, Genetics and Molecular Biology</term>
         <term level="3">Genetics</term>
       </keywords>
      ▼<keywords change="#rd-multicat" resp="#istex" scheme="#scopus">
         <term level="1">Life Sciences</term>
         <term level="2">Biochemistry, Genetics and Molecular Biology</term>
         <term level="3">Biochemistry</term>
       </keywords>
      ▼<keywords change="#rd-multicat" resp="#istex" scheme="#scopus">
         <term level="1">Life Sciences</term>
         <term level="2">Biochemistry, Genetics and Molecular Biology</term>
         <term level="3">Structural Biology</term>
       </keywords>
     </annotationBlock>
   </ns1:listAnnotation>
  </ns1:standOff>
</TEI>
```

Exemple d'extraction d'entités nommées (corps du texte)

Enrichments

unitex TEI

```
<TEI>
    <teiHeader>
        …
    </teiHeader>
    <text>
        …
    </text>
    <standOff>
        …
        …
    </standOff>
</TEI>
```

annotations

Revue internationale de sémiotique juridique Vol.VI no.18 [1993]

TROIS FIGURES D'INTERACTI

Compte-rendu de *Lire le droit. Langue, texte, cognit*
de D. Bourcier et P. Mackay, Paris, L.G.D.J., C
1992, 486 pp., ISBN 2-275-00567-6, 260 FF.

Il est des séminaires auxquels on regrette vivement
lorsque l'on en vient à lire les communications qui y fu
le cas de celui tenu à l'abbaye de Royaumont du 28 au
la direction de Danièle Bourcier et Pierre Mackay, et q
spécialistes de disciplines cognitives. Le résultat éc
fourni, divisé en sept chapitres, qui envisage le droit
d'un point de vue cognitif. Ce dernier sera entendu
Danièle Bourcier qui rappelle, dans sa présentation, c
tives sont définies "comme l'étude interdisciplinaire
représentation et du traitement des connaissances."
disciplines ne doit pas, selon elle, faire oublier qu' "ell
un objet commun, le texte" (12).

```
<TEI xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns="http://www.tei-c.org/ns/1.0"
    xsi:noNamespaceSchemaLocation="https://xml-schema.delivery.istex.fr/formats/tei-istex.xsd"
    xmlns:nsl="http://standoff.proposal">
    <nsl:standOff>
        <teiHeader> [43 lines]
        <nsl:listAnnotation type="orgName" xml:lang="fr"> [8 lines]
        <nsl:listAnnotation type="date" xml:lang="fr"> [15 lines]
        <nsl:listAnnotation type="ref" subtype="bibl" xml:lang="fr"> [8 lines]
        <nsl:listAnnotation type="placeName" xml:lang="fr">
            <annotationBlock corresp="text">
                <placeName change="#Unitex-3.2.0-alpha" resp="istex-rd"
                    scheme="https://placename-entity.data.istex.fr">
                    <term>abbaye de Royaumont</term>
                    <fs type="statistics">
                        <f name="frequency">
                            <numeric value="1"/>
                        </f>
                    </fs></placeName>
            </annotationBlock>
            <annotationBlock corresp="text">
                <placeName change="#Unitex-3.2.0-alpha" resp="istex-rd"
                    scheme="https://placename-entity.data.istex.fr">
                    <term>Paris</term>
                    <fs type="statistics">
                        <f name="frequency">
                            <numeric value="5"/>
                        </f>
                    </fs>
                </placeName>
            </annotationBlock>
        </nsl:listAnnotation>
        <nsl:listAnnotation type="persName" xml:lang="fr"> [92 lines]
    </nsl:standOff>
</TEI>
```

Fréquence d'apparition dans le texte

Conclusion

ISTEX > En Bref

Article Journal *Nature*
https://api.istex.fr/ark:/67375/GT4-LHM354SF-Z/fulltext.pdf?sid=istex-api-demo

A novel protein–mineral interface

Transferrins transport Fe3+ and other metal ions in mononuclear-binding sites. We present the first evidence that a member of the transferrin superfamily is able to recognize multi-nuclear oxo-metal clusters, small mineral fragments that are the most abundant forms of many metals in the environment. We show that the ferric…

nature | research-article
ark:/67375/GT4-LHM354SF-Z
Score : 9.88
Mots : 4880
Publication : 2003

Fulltext | Metadata | Covers | Annexes | Enrichments

PDF ZIP XML GIF GIF

Fichiers source éditeur

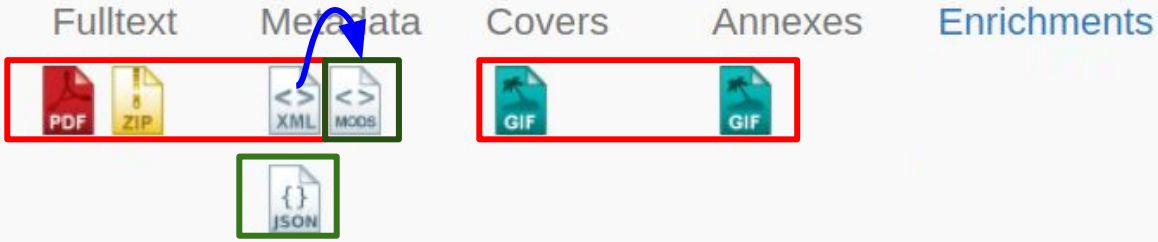
ISTEX > En Bref

Article Journal *Nature*
https://api.istex.fr/ark:/67375/GT4-LHM354SF-Z/fulltext.pdf?sid=istex-api-demo

A novel protein–mineral interface

Transferrins transport Fe3+ and other metal ions in mononuclear-binding sites. We present the first evidence that a member of the transferrin superfamily is able to recognize multi-nuclear oxo-metal clusters, small mineral fragments that are the most abundant forms of many metals in the environment. We show that the ferric…

nature | research-article
ark:/67375/GT4-LHM354SF-Z
Score : 9.88
Mots : 4880
Publication : 2003



Fichiers source éditeur

Reformatage format Pivot

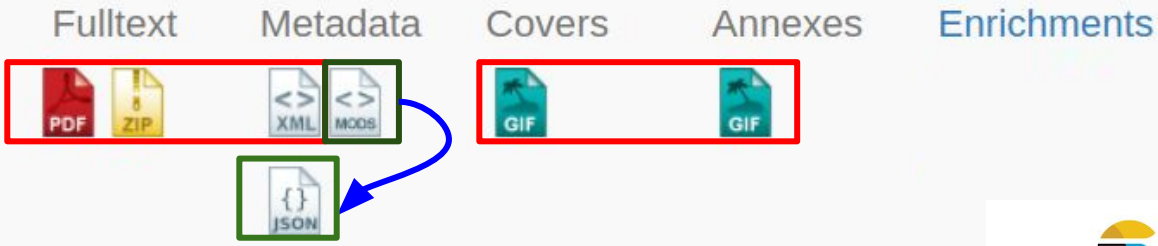
ISTEX > En Bref

Article Journal *Nature*
https://api.istex.fr/ark:/67375/GT4-LHM354SF-Z/fulltext.pdf?sid=istex-api-demo

A novel protein–mineral interface

Transferrins transport Fe3+ and other metal ions in mononuclear-binding sites. We present the first evidence that a member of the transferrin superfamily is able to recognize multi-nuclear oxo-metal clusters, small mineral fragments that are the most abundant forms of many metals in the environment. We show that the ferric…

nature | research-article
ark:/67375/GT4-LHM354SF-Z
Score : 9.88
Mots : 4880
Publication : 2003



Indexation moteur de recherche

elasticsearch

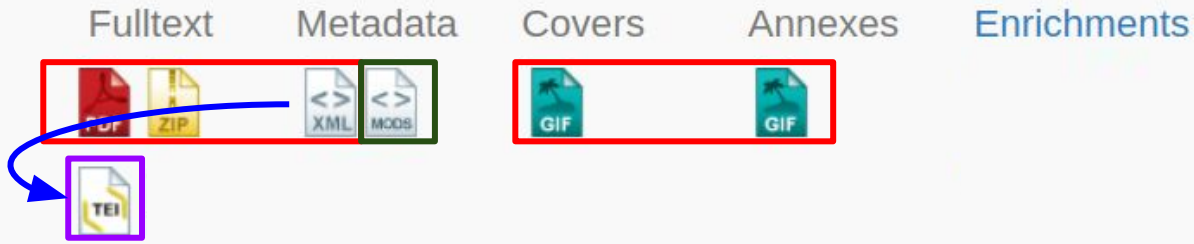
ISTEX > En Bref

Article Journal *Nature*
https://api.istex.fr/ark:/67375/GT4-LHM354SF-Z/fulltext.pdf?sid=istex-api-demo

A novel protein–mineral interface

Transferrins transport Fe3+ and other metal ions in mononuclear-binding sites. We present the first evidence that a member of the transferrin superfamily is able to recognize multi-nuclear oxo-metal clusters, small mineral fragments that are the most abundant forms of many metals in the environment. We show that the ferric…

nature | research-article
ark:/67375/GT4-LHM354SF-Z
Score : 9.88
Mots : 4880
Publication : 2003



Fichiers source éditeur

Reformatage en TEI

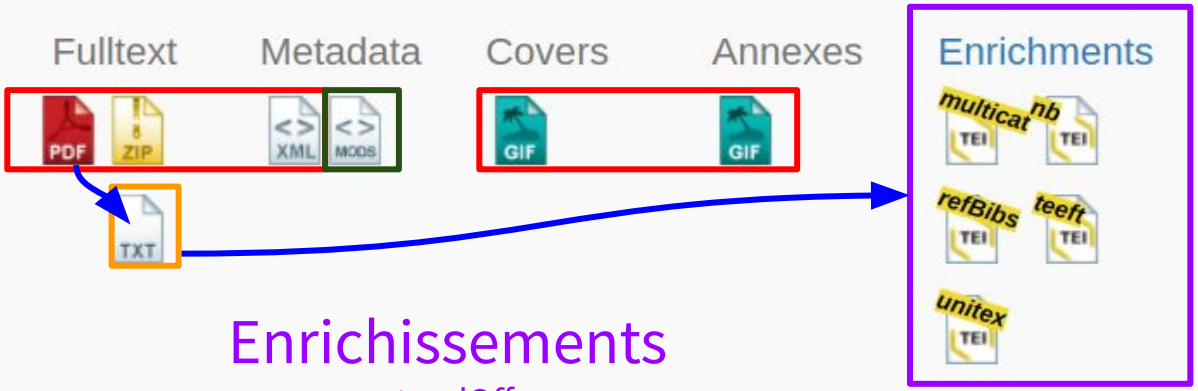
ISTEX > En Bref

Article Journal *Nature*
https://api.istex.fr/ark:/67375/GT4-LHM354SF-Z/fulltext.pdf?sid=istex-api-demo

A novel protein–mineral interface

Transferrins transport Fe3+ and other metal ions in mononuclear-binding sites. We present the first evidence that a member of the transferrin superfamily is able to recognize multi-nuclear oxo-metal clusters, small mineral fragments that are the most abundant forms of many metals in the environment. We show that the ferric…

nature | research-article
ark:/67375/GT4-LHM354SF-Z
Score : 9.88
Mots : 4880
Publication : 2003



Enrichissements
<standOff>

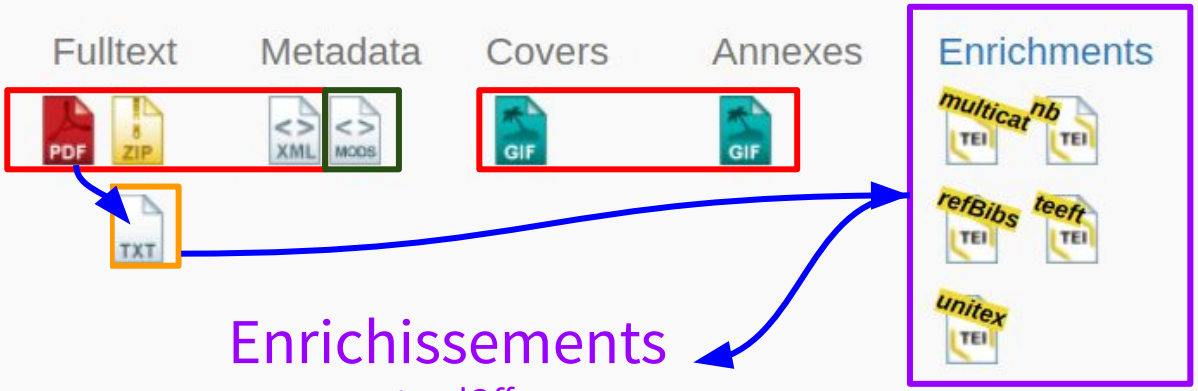
ISTEX > En Bref

Article Journal *Nature*
https://api.istex.fr/ark:/67375/GT4-LHM354SF-Z/fulltext.pdf?sid=istex-api-demo

A novel protein–mineral interface

Transferrins transport Fe3+ and other metal ions in mononuclear-binding sites. We present the first evidence that a member of the transferrin superfamily is able to recognize multi-nuclear oxo-metal clusters, small mineral fragments that are the most abundant forms of many metals in the environment. We show that the ferric…

nature | research-article
ark:/67375/GT4-LHM354SF-Z
Score : 9.88
Mots : 4880
Publication : 2003



Enrichissements
<standOff>

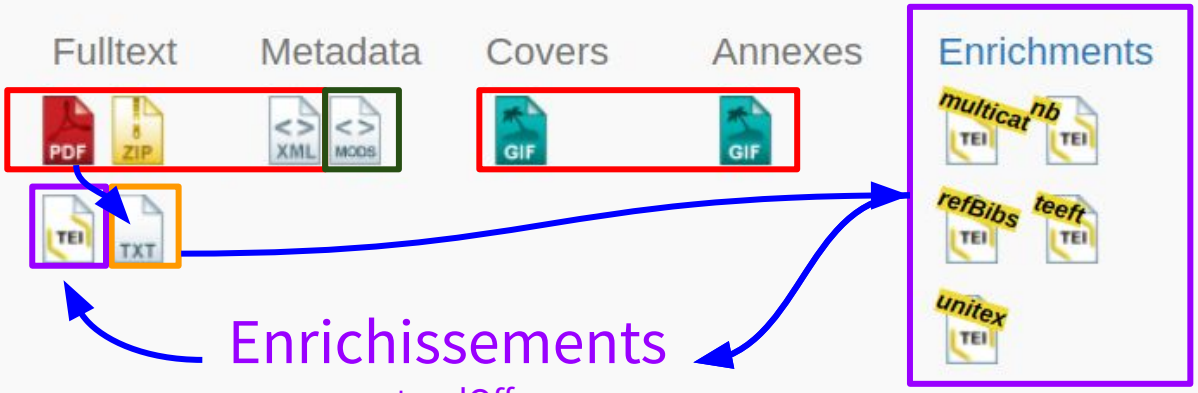
ISTEX > En Bref

Article Journal *Nature*
https://api.istex.fr/ark:/67375/GT4-LHM354SF-Z/fulltext.pdf?sid=istex-api-demo

A novel protein–mineral interface

Transferrins transport Fe3+ and other metal ions in mononuclear-binding sites. We present the first evidence that a member of the transferrin superfamily is able to recognize multi-nuclear oxo-metal clusters, small mineral fragments that are the most abundant forms of many metals in the environment. We show that the ferric…

nature | research-article
ark:/67375/GT4-LHM354SF-Z
Score : 9.88
Mots : 4880
Publication : 2003



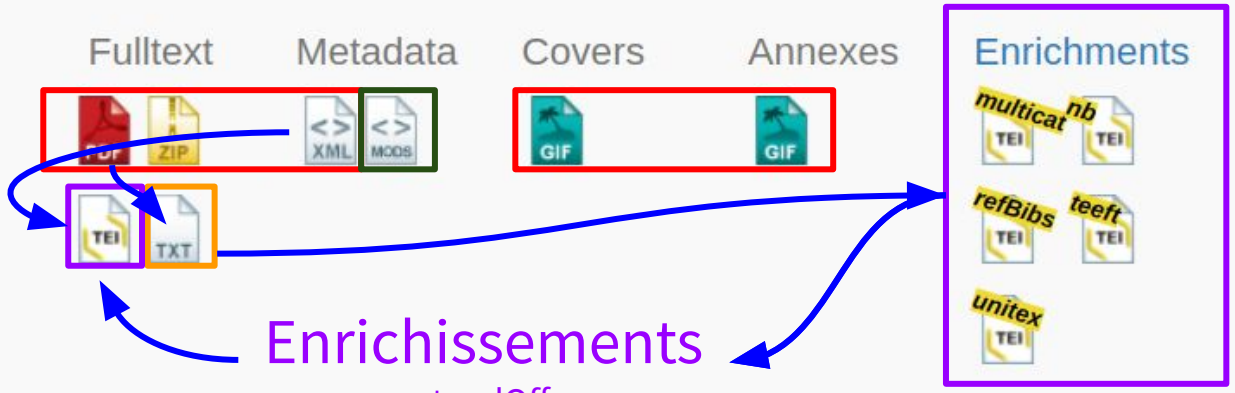
ISTEX > En Bref

Article Journal *Nature*
https://api.istex.fr/ark:/67375/GT4-LHM354SF-Z/fulltext.pdf?sid=istex-api-demo

A novel protein–mineral interface

Transferrins transport Fe3+ and other metal ions in mononuclear-binding sites. We present the first evidence that a member of the transferrin superfamily is able to recognize multi-nuclear oxo-metal clusters, small mineral fragments that are the most abundant forms of many metals in the environment. We show that the ferric…

nature | research-article
ark:/67375/GT4-LHM354SF-Z
Score : 9.88
Mots : 4880
Publication : 2003



ISTEX > TEI en route vers la fouille de texte

Article Journal *Nature*
https://api.istex.fr/ark:/67375/GT4-LHM354SF-Z/fulltext.pdf?sid=istex-api-demo

A novel protein–mineral interface

Transferrins transport Fe3+ and other metal ions in mononuclear-binding sites. We present the first evidence that a member of the transferrin superfamily is able to recognize multi-nuclear oxo-metal clusters, small mineral fragments that are the most abundant forms of many metals in the environment. We show that the ferric…

nature | research-article
ark:/67375/GT4-LHM354SF-Z
Score : 9.88
Mots : 4880
Publication : 2003

Fulltext | Metadata | Covers | Annexes | Enrichments

PDF ZIP XML MODS GIF GIF multicat nb refBibs teeft unitex
TEI TXT JSON

Source: https://www.quentinfily.fr/wp-content/uploads/2015/11/fouille-de-texte-300x225.jpg

S'authentifier :

- Vérifier ses droits d'accès : https://api.istex.fr/auth

Se documenter :

- ODD-istex: https://tei.istex.fr/
- Documentation "Usage TDM d'ISTEX" : https://doc.istex.fr/tdm/
- Documentation "API ISTEX" :

Se connecter :

- ISTEX : http://www.istex.fr
- Démonstrateur ISTEX : http://demo.istex.fr/

Merci

Chercher de l'aide / Contribuer à l'amélioration :

- Contact :
  - Via le formulaire : https://www.istex.fr/contact/
  - Via la liste : contact@listes.istex.fr
- Liste publique de discussion : users@listes.istex.fr

Se tenir informé :

- Blog ISTEX : https://blog.istex.fr/
- Plateforme Twitter : @ISTEX_Platform