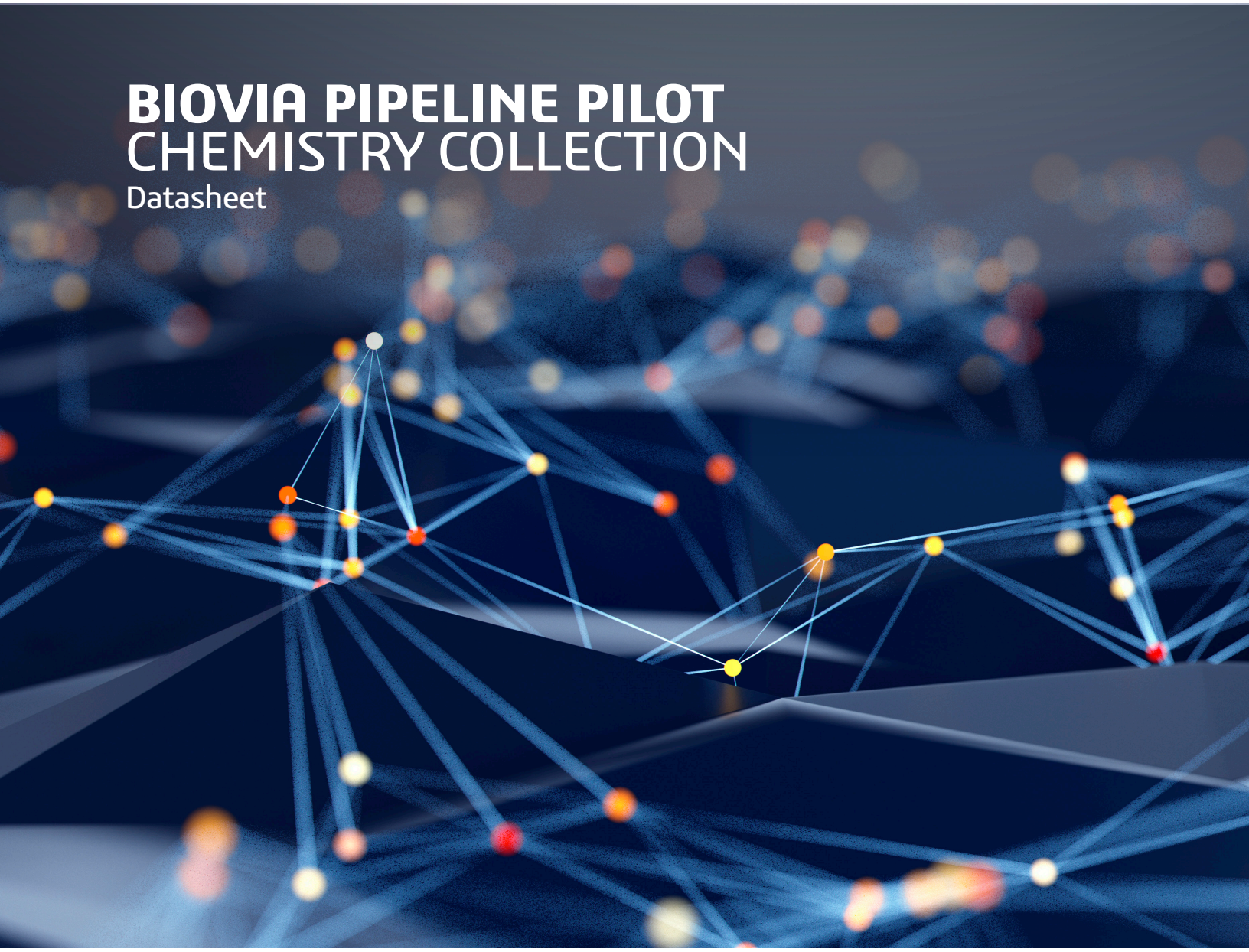
BIOVIA

3DEXPERIENCE®

BIOVIA PIPELINE PILOT CHEMISTRY COLLECTION

Datasheet



The Chemistry Collection offers a comprehensive suite of readers, writers, viewers, molecular property calculators, filters, and manipulators. This collection of modular components extends the standard capabilities of BIOVIA Pipeline Pilot to include compound processing and cheminformatics analyses. With these enhancements, you can create protocols for a broad array of chemistry applications.

DASSAULT SYSTEMES

CHEMISTRY COLLECTION COMPONENTS

Readers and Writers

The Chemistry Collection includes file readers and writers for most industry-standard chemistry formats as listed below:

- BIOVIA formats MOL file, SD, SDF, RG, SKC , XD, XDF, RXN, RD, RDF (BIOVIA Formats via MDL);
- Pistoia Alliance formats UDM, HELM 1 and 2, XHELM;
- Other formats such as ACCORD, SMILES, SMIRKS, SMARTS, InChI, RInChI, OpenEye OEMeta Data (name to structure, structure to name), MOL2 (from Tripos), CDX, CDXML (from Perkin-Elmer); and public formats such as PDB.

Viewers

In addition to BIOVIA Pipeline Pilot's standard viewers, the Chemistry Collection seamlessly integrates several popular structural viewers including BIOVIA Discovery Studio Visualizer and BIOVIA Insight for Excel. Furthermore, a variety of Web viewers are available providing a way to graphically display your molecular data.

Manipulators

The Chemistry Collection includes manipulators that modify structures according to their salts, tautomers, stereochemistry, charges, and more. It allows to normalize sets of molecules before comparing them, apply chemical reactions, merge collections of molecules and even enumerate combinatorial libraries. For optimal display characteristics, the user can employ manipulators for structural alignment and for fast, high quality 2D and 3D layout.

Filters

Numerous filter components are provided to separate molecules into groups based on chemical properties or calculations. These filters range from simple property value thresholds, similar to those in the Lipinski Filter, to more sophisticated criteria, such as occurrences of substructures or duplicates. Filters are available to select representative subsets based on the diversity or similarity. An interactive filter allows selection a subset of molecules for further processing. What's more is the ability to create custom chemistry filters for any property values that you require.

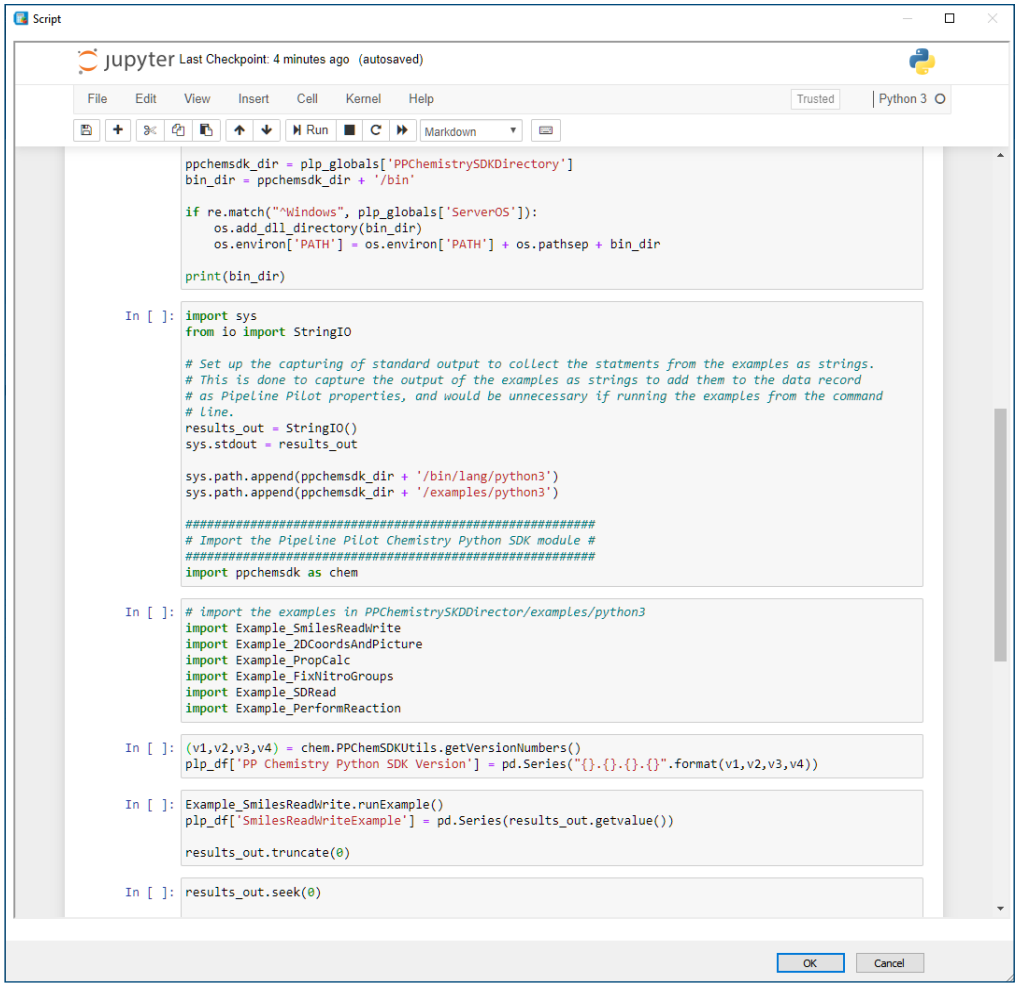


Figure 1. Example of Pipeline Pilot Chemistry SDK Python code written with Jupyter Notebook

Property Calculators

The Chemistry Collection includes chirality calculation (CIP – Cahn, Ingold, Prelog), a variety of fast molecular property calculators. You can characterize hundreds or thousands of molecules per second. Whole molecule properties include AlogP, logD, pKa and solubility, and molecular weight. Other properties include Synthetic Accessibility Score (SA Score), QED (Drug Likeness), solubility, Rule of 5 (Lipinski), electrotopological indices, solvent accessible surface area and volume descriptors and a large variety of molecular property counts.

Structural Fingerprints

The Chemistry Collection includes a proprietary method for calculating structural fingerprints. Known as Extended Connectivity Fingerprints (ECFP), this method offers excellent characterization of molecules, indexing the environments of every atom in a molecule using up to four billion different structural features. It is an extremely rapid method useful in searching, clustering, and modeling applications. When used with Bayesian learning technologies in the separately available Modeling Collection, ECFP provides interpretable predictive models for large data sets. You can also use Daylight, MDL, or user defined keys within the system.

BENEFITS OF BIOVIA PIPELINE PILOT CHEMISTRY COLLECTION:

Enumerate Compounds

- Matched Molecular Pairs analysis
- Ring Assemblies Replacement, Scaffold Replacement (eg, Murcko)
- Chain Replacement
- Rgroup, stereo, tautomer, reaction enumeration

Manage Compound Datasets

- Clean up and standardize your datasets compare multiple datasets
- Search for substructure and similarity across datasets
- Perform extensive property profiling and subset selection

Build and Manage Databases with Direct Cartridge Components

- Import data to build and load databases
- Store molecules, reactions and CMBs
- Build and manage chemistry index, search within built databases

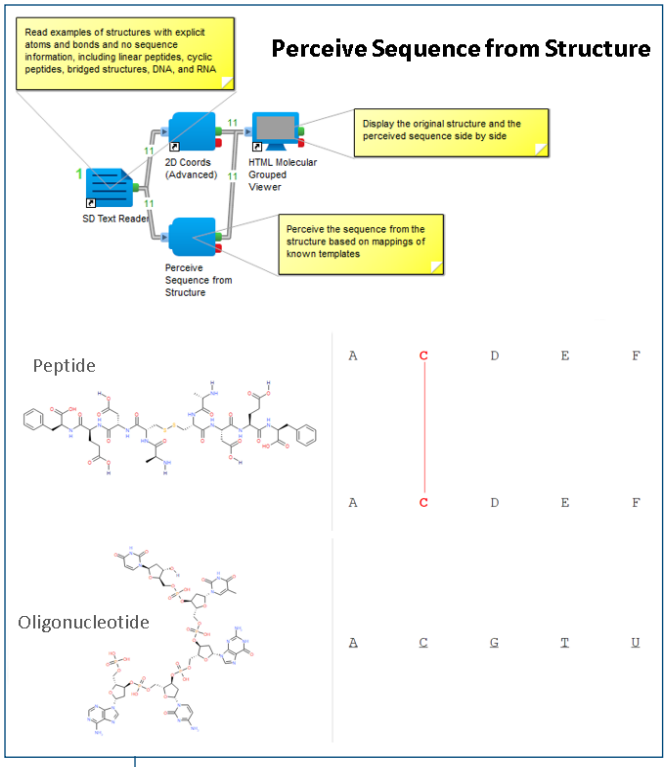


Figure 2. Generate Self Contained Sequence Representation (SCSR) molfiles from older molfile formats using the Perceive Sequence protocol

Access Molecular Toolkit

- Build custom chemistry components with the Molecular Toolkit's Pilot Script, Python, Java and Perl, to use in Pipeline Pilot protocols and to implement custom functionality that isn't included with the standard Pipeline Pilot Chemistry Collection

Access Pipeline Pilot Chemistry SDK

- Build custom applications without a Pipeline Pilot Server but with all the chemistry features using the Pipeline Pilot Chemistry SDK's supported languages: Python (use Jupyter Notebook to simplify code development), Java and .NET

Work with Chemistry Sketchers

- Use BIOVIA Pipette Sketcher, BIOVIA Draw or other sketchers such as Revvity's ChemDraw to sketch your chemistry entities to use in your components in the Chemistry Collection.

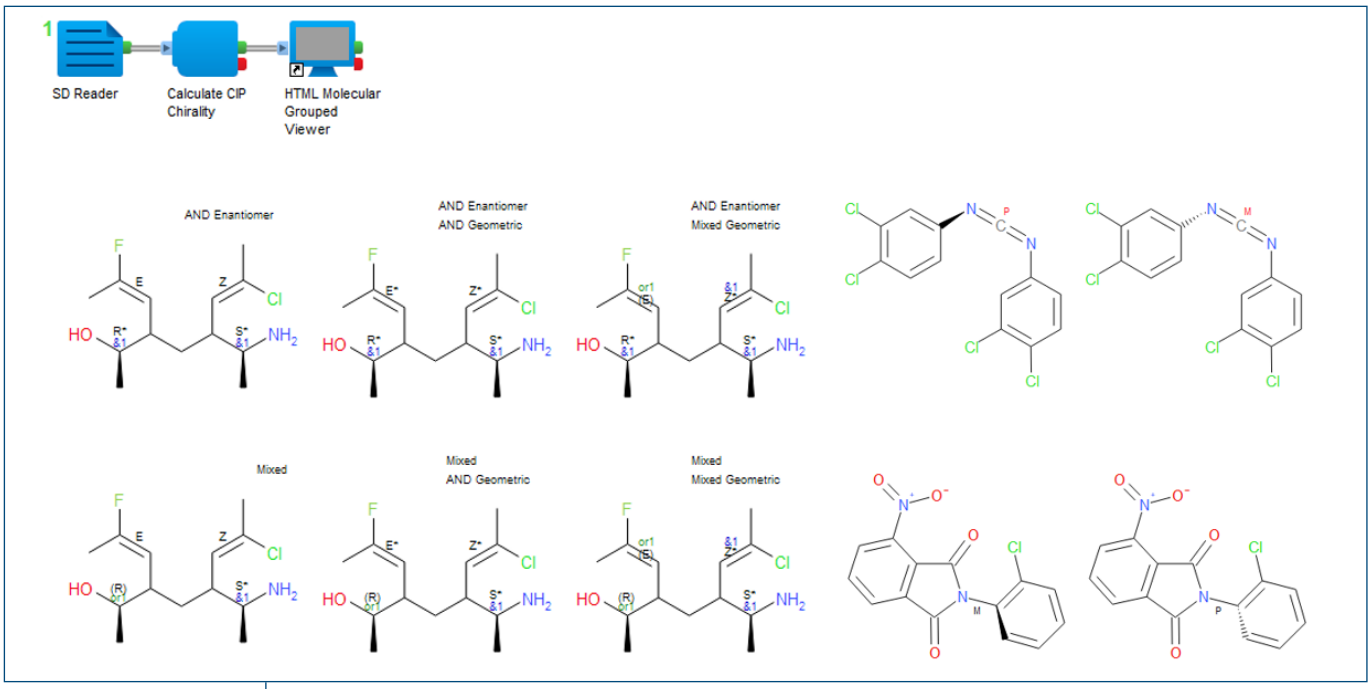


Figure 3. Example of the new enhanced Geometric stereoisomers and other stereochistry depiction options

Work with Chemically Modified Biomolecules

The Chemistry Collection includes a set of components and protocols to help scientists to read, write, display sequences and chemically modified biomolecules (peptides, recombinant proteins, ADCs, Oligonucleotides)

- Perceive sequences from older formats
- Interconvert between BIOVIA and other formats (eg, HELM, PDB)
- Create a centralized library: Manage Monomers and SCSR templates used to represent chemically modified biomolecules (CMB)

Utilize Comprehensive Stereochemistry Support

The Chemistry Collection and the Chemistry SDK support the ability to Read, Write, Perceive, Depict and Search for the following classes of stereo and geometric isomers:

- Tetrahedral including meso, pseudo, anti/syn
- Axial for Allenes and Atropisomers
- Higher order Stereochemistry: Octahedral, bipyramidal, square planner
- Geometric – cis/trans double bonds
- Enhanced stereochemistry for atoms and double bonds – relative stereogroups: ABS, AND and OR

BENEFITS OF PIPELINE PILOT CHEMISTRY COLLECTION, COMBINED WITH THE SEPARATELY AVAILABLE MODELING COLLECTION:

- Perform structure activity modeling
- Cluster compounds
- Search for maximal common substructure

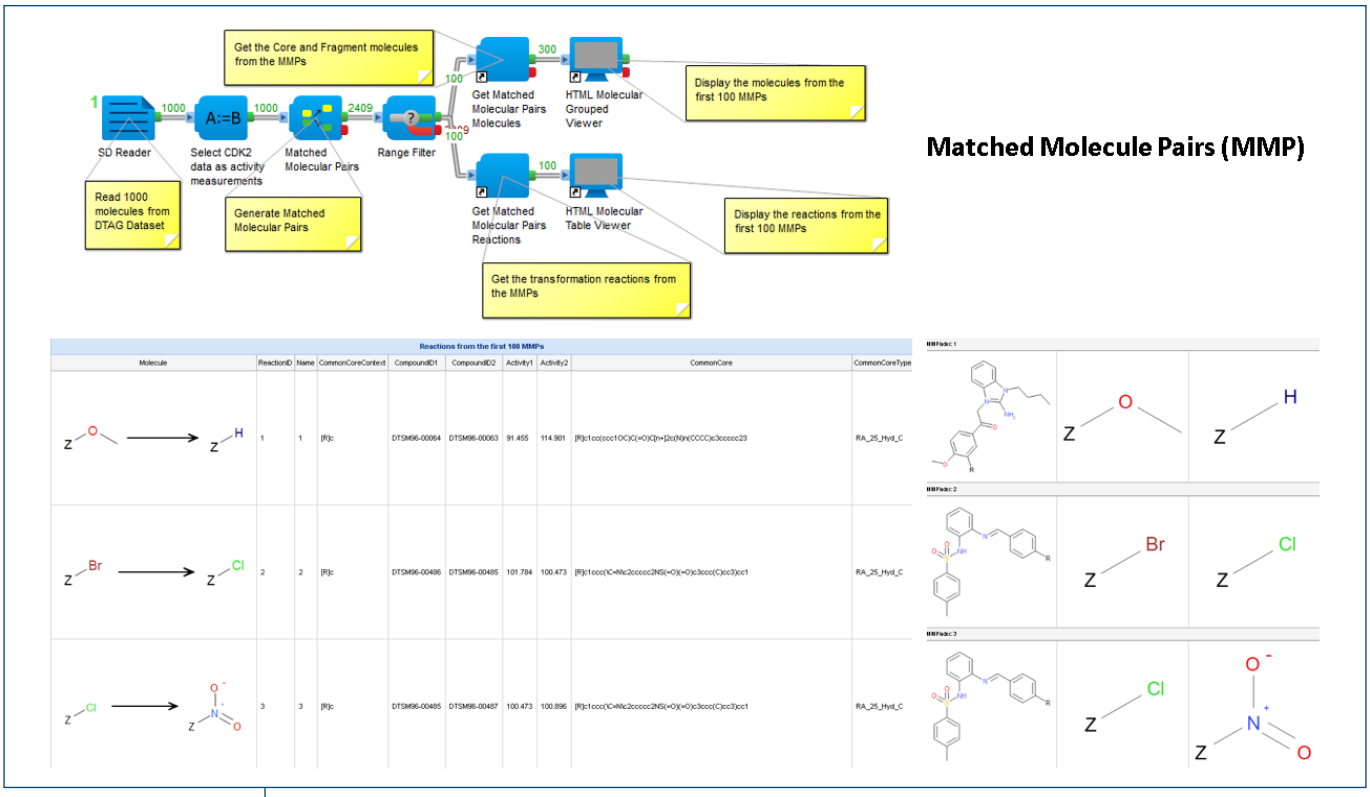


Figure 4. Example protocol to generate a MMP (Matched Molecule Pairs) for a set of compounds

Our **3D**EXPERIENCE® platform powers our brand applications, serving 11 industries, and provides a rich portfolio of industry solution experiences.

Dassault Systèmes, the **3DEXPERIENCE** Company, is a catalyst for human progress. We provide business and people with collaborative virtual environments to imagine sustainable innovations. By creating 'virtual experience twins' of the real world with our **3DEXPERIENCE** platform and applications, our customers push the boundaries of innovation, learning and production.

Dassault Systèmes' 20,000 employees are bringing value to more than 270,000 customers of all sizes, in all industries, in more than 140 countries. For more information, visit **www.3ds.com**.

3DEXPERIENCE®

DASSAULT SYSTEMES | The **3DEXPERIENCE**® Company

Americas
Dassault Systèmes
175 Wyman Street
Waltham, Massachusetts
02451-1223
USA

Europe/Middle East/Africa
Dassault Systèmes
10, rue Marcel Dassault
CS 40501
78946 Vélizy-Villacoublay Cedex
France

Asia-Pacific
Dassault Systèmes K.K.
ThinkPark Tower
2-1-1 Osaki, Shinagawa-ku,
Tokyo 141-6020
Japan