Vyhledávání podstruktur je jednou z nejcennějších schopností databází malých molekul. Dostupné databáze typicky poskytují akceptovatelně rychlé zpracování uživatelských dotazů, ale nejsou dostatečně škálovatelné s velikostí uložených dat. V této práci je popsána nová open-source databáze Sachem, která implementuje novoý způsob vyhledávání podstruktur využívající nově sestavené otisky chemických molekul uložené v invertovaných databázových indexech. Rychlost vyhledávání v této databázi byla měřena na datových sadách obsahujících desítky milionů molekul. Porovnání výkonnosti s jinými dostupnými databázemi potvrdilo zlepšení v celkové rychlosti hledání, možností škálování výkonnosti i v efektivitě prosívání dat. Práce dále popisuje aplikaci databáze Sachem, službu založenou na dotazovacím jazyku SPARQL, která rozšiřuje existující sémantické datové služby o možnost zahrnout v dotazech i chemicky relevantní strukturní a podobnostní podmínky. Výsledek nabízí nové, jednodušší možnosti dotazování v dostupných heterogenních datových zdrojích.