

## LA CRISTALLOGRAFIA NELLE SCIENZE DELLA VITA: SIAMO VICINI AL LIMITE?

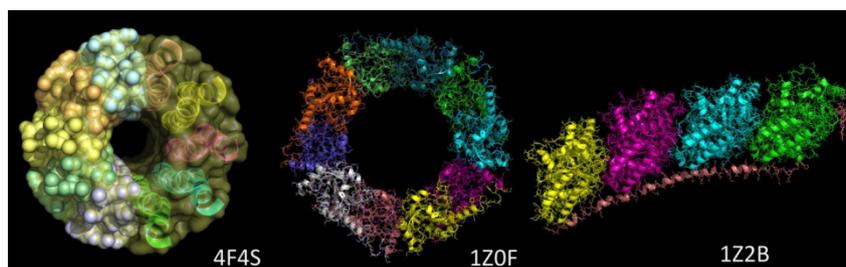
Giuseppe Zanotti

Dipartimento di Scienze Biomediche

Università di Padova

[giuseppe.zanotti@unipd.it](mailto:giuseppe.zanotti@unipd.it)

<http://tiresia.bio.unipd.it/zanotti>



Rappresentazione schematica di tre strutture macromolecolari altamente simmetriche. Da sinistra a destra:  $F_1F_0$  ATPase da lievito (PDB code 4F4S), Alkyl Hydroperoxide-reductase da *Helicobacter pylori* (PDB code 1Z0F) e complesso tubulin-colchicine-vinblastine (PDB code 1Z2B). Le immagini delle strutture sono state prodotte utilizzando il software Pymol (The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC)

La cristallografia macromolecolare ha vissuto straordinari progressi dagli anni Cinquanta, quando sono state determinate le prime strutture tridimensionali di proteine, fino a oggi. Il continuo aumento della potenza di calcolo, la disponibilità di nuove sorgenti di raggi X quali i sincrotroni, la possibilità di esprimere le proteine in sistemi eterologhi, i nuovi e sempre più potenti algoritmi di calcolo hanno reso la determinazione e l'affinamento delle strutture macromolecolari quasi un lavoro di routine. Eppure, le nuove metodiche basate sui laser a raggi X a elettroni liberi stanno per introdurre nel settore una nuova e forse definitiva rivoluzione.

### Crystallography in Life Sciences: Are We close to the Limit?

Macromolecular crystallography has witnessed extraordinary advancements since the end of 1950s, when the first protein structures were determined. The improvement of computer power, the availability of powerful X-ray sources such as synchrotrons, the possibility of expressing proteins in heterologous systems, and new algorithms for structure determination and refinement have already made the determination of the three-dimensional structure of a macromolecule quite straightforward. Nevertheless, the emergence of a new methodology based on x-ray free-electron lasers promises to introduce another revolution in the field.

### La prospettiva storica

Dopo il primo esperimento di diffrazione dei raggi X effettuato da von Laue, Friederich e Knipping nel 1912<sup>1</sup>, la ricerca cristallografica nella prima metà del XIX secolo si era concentrata soprattutto sullo sviluppo delle metodologie appropriate per misurare i dati di diffrazione e per risolvere la struttura da cristallo singolo. L'assenza di calcolatori elettronici, per il cui utilizzo si è dovuto aspettare la fine della seconda guerra mondiale, limitava fortemente la possibilità di affrontare i lunghi e tediosi calcoli necessari per portare a termine le trasformate di Fourier tridimensionali; le strutture risolte nella prima metà del secolo sono state essenzialmente

quelle di cristalli ionici con pochi atomi nella cella cristallina o di cristalli molecolari contenenti molecole con un numero molto limitato di atomi. Per le prime strutture tridimensionali di proteine dobbiamo aspettare la fine degli anni Cinquanta, con la proteina che lega l'ossigeno, la mioglobina, pubblicata nel 1958 da Kendrew e collaboratori<sup>2</sup>, seguita subito dopo da quella del trasportatore dell'ossigeno nel plasma, l'emoglobina, da parte di Max Perutz e collaboratori<sup>3</sup>. Il campo della cristallografia di proteine (il termine "cristallografia di proteine" può essere fuorviante e dovrebbe essere sostituito da "cristallografia di macromolecole", poiché la tecnica non si limita allo studio della struttura delle proteine, ma include tutte le altre macromolecole biologiche, quali RNA, DNA, polisaccaridi, virus ecc.; tuttavia, poiché la stragrande maggioranza delle strutture risolte è rappresentata dalle proteine globulari, la prima espressione viene comunemente usata) procedette abbastanza lentamente durante i primi vent'anni. Fino alla fine degli anni Settanta la determinazione di una struttura completamente nuova poteva ancora prendere da cinque a dieci anni di intenso lavoro!

Nel 1972, anno della sua fondazione, solo dodici strutture di proteine erano presenti nella "Protein Data Bank" (PDB), la banca dati che raccoglie le coordinate e i dati di diffrazione di tutte le strutture macromolecolari pubblicate nel mondo. Oggi<sup>a</sup> vi si trovano più di 80.000 strutture determinate mediante diffrazione dei raggi X, liberamente accessibili ([www.pdb.org](http://www.pdb.org)). Il grafico di Fig. 1 permette di apprezzare la crescita pressoché esponenziale, iniziata negli anni Novanta, del numero di strutture depositate. Questa crescita è tuttora in atto, sebbene qualche diminuzione della pendenza stia forse diventando visibile, e le sue ragioni possono essere fatte risalire a quattro fattori principali: *i*) la continua crescita della potenza di calcolo disponibile, *ii*) il miglioramento continuo degli algoritmi e dei programmi usati per risolvere il problema della fase, *iii*) i progressi della biologia molecolare e *iv*) la disponibilità delle sorgenti di luce di sincrotrone. Il primo punto non necessita di spiegazioni: ciascuno di noi ha sperimentato il fatto che oggi con un personal computer si possono effettuare calcoli che negli anni Ottanta richiedevano un supercomputer. Il secondo punto è molto tecnico e le persone interessate possono trovare interi libri sull'argomento<sup>4</sup>. La biologia molecolare ha reso disponibili per la cristallizzazione proteine che sono presenti solo in tracce nella cellula, o da sorgenti che sono difficilmente reperibili (si consideri, per esempio, che se si vuole studiare una proteina di origine umana, solo le proteine del plasma sono facilmente ottenibili). Infine, i vantaggi dell'uso della radiazione di sincrotrone sono commentati nel prossimo paragrafo. Nei successivi si discute delle prospettive e dei limiti della tecnica.

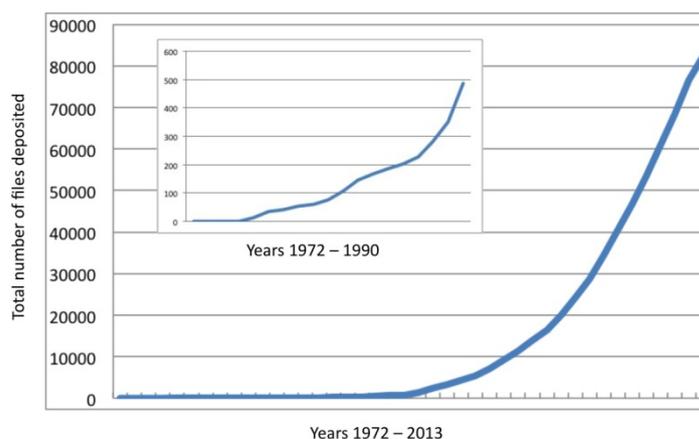


Fig. 1 - La crescita della Protein data Bank. Il grafico illustra l'aumento del numero totale di strutture depositate dal 1972, anno di fondazione della banca dati. Nell'inserito, si vede il dettaglio della crescita dal 1972 al 1990

<sup>a</sup>Alla data del 20 agosto 2013, erano presenti nella PDB 93.252 strutture, delle quali 82.337 determinate per diffrazione dei raggi X, 10.051 per NMR e le altre per microscopia elettronica o tecniche miste.

### La radiazione di sincrotrone

È difficile sottostimare l'importanza che i sincrotroni hanno assunto nella cristallografia di macromolecole (e naturalmente in molti altri campi, al di fuori degli scopi di quest'articolo). Le apparecchiature di terza generazione forniscono un fascio di raggi X estremamente brillante (la brillantezza del fascio, definita come il numero di fotoni/secondo/mm<sup>2</sup>/0,1% di banda passante, può essere di dieci-dodici ordini di grandezza o più rispetto a quella di un generatore a tubo sigillato) e, su linee equipaggiate con rivelatori adatti, la misura di un intero set di dati di diffrazione può essere effettuata in due-tre minuti; si consideri che nei vecchi tempi, usando una camera a rotazione, erano necessarie almeno ventiquattro ore solo per impressionare le pellicole, e anche più tempo poi per processare i dati. Inoltre, sono oggi operative linee micro-focus sulle quali si possono misurare i dati di diffrazione da cristalli singoli di dimensioni molto piccole, fino a qualche micrometro per lato.

Un altro importante vantaggio della radiazione di sincrotrone rispetto a generatori convenzionali è quella che si definisce *tunability*, cioè la possibilità di selezionare specifiche lunghezze d'onda. Quest'ultima proprietà permette di risolvere il problema della fase usando la tecnica della dispersione o diffrazione anomala. Poiché, infatti, con le moderne tecniche di biologia molecolare si possono produrre proteine nelle quali l'amminoacido metionina è sostituito da selenometionina, la misura dei dati di diffrazione alla lunghezza d'onda opportuna può essere usata per risolvere il problema della fase e determinare la struttura tridimensionale, in modo molto spesso quasi automatico.

### Oggetti molto grandi

Esiste un limite alle dimensioni oltre il quale le macromolecole o i complessi macromolecolari possono essere studiati usando la diffrazione dei raggi X? La risposta è no. La sola condizione limitante è che l'oggetto che si vuole studiare si possa ottenere nella forma di un cristallo singolo, grande a sufficienza per la diffrazione (per quanto riguarda il significato del termine "cristallo grande", si veda il paragrafo su XFEL).

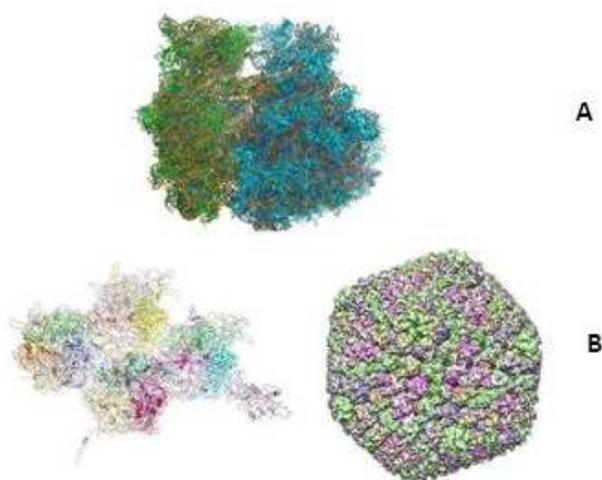
Naturalmente, più grande è la molecola sotto indagine, più laboriosa è la determinazione della sua struttura. In anni recenti abbiamo assistito alla realizzazione di un progetto in corso da molti anni, la determinazione della struttura del ribosoma. Quest'ultimo è il marchingegno molecolare che assiste, catalizza e coordina la biosintesi delle proteine partendo dal suo RNA messaggero. Per il raggiungimento di questo traguardo, nel 2009 V. Ramakrishnan, T.A. Steitz e A.E. Yonath hanno vinto il premio Nobel per la chimica. Il ribosoma è un complesso, trascurando gli atomi di idrogeno, di circa 60.000 atomi (Fig. 2A). La sua struttura si è rivelata fondamentale per chiarire il meccanismo molecolare che controlla la produzione di proteine nella cellula e, grazie alla struttura di complessi ribosoma/inibitore, ha permesso e permetterà di disegnare nuovi e più potenti antibiotici<sup>5</sup>.

Un'altra classe di oggetti molto grandi studiati dalla cristallografia sono i virus. I virus non sono molecole e neppure complessi macromolecolari: essi sono formati da un guscio proteico, il capsido, che include al proprio interno l'acido nucleico, DNA o RNA. I virus sono una via di mezzo tra la materia non-vivente e vivente, poiché non sono in grado di riprodursi da soli e necessitano, per farlo, di una cellula ospite. Il capsido dei virus ha un'elevata simmetria e quest'ultima ne favorisce la cristallizzazione<sup>b</sup>. Alla data di ottobre 2013 erano presenti nella PDB 287 file di coordinate atomiche di virus sferici ottenute tramite diffrazione dei raggi X<sup>c</sup>. I più grandi tra questi sono l'adenovirus umano<sup>6</sup> e il virus adeno-associato sierotipo 3B<sup>7</sup>, che contano rispettivamente 92.000 e 83.000 atomi nell'unità asimmetrica. La struttura dell'adenovirus è illustrata in Fig. 2B.

---

<sup>b</sup>Per essere più precisi, solo i virus sferici privi di membrana esterna cristallizzano.

<sup>c</sup>Altre 60 strutture di virus erano state determinate per microscopia elettronica.



*Fig. 2 - Struttura del ribosoma 70 S da *Thermus thermophilus*. La subunità 30S è colorata in verde, la 50 in azzurro (coordinate PDB 2WDK e 2WDL). Il tracciato della catena di RNA ribosomiale è mostrato come un sottile tubo arancio. Fig. 2A: vista dell'unità asimmetrica dell'adenovirus umano (sinistra, coordinate PDB 1VSZ) e rappresentazione schematica dell'intero capsid (destra). La Fig. 2B è riprodotta dalla PDB*

### **Luci e ombre. Il lato oscuro**

Il maggiore ostacolo alla biologia strutturale basata sui raggi X è rappresentato dalla necessità, per poter determinare la struttura di una macromolecola, di ottenerne cristalli di sufficiente grandezza e qualità. La procedura di cristallizzazione rappresenta oggi lo stadio limitante, mentre gli altri passaggi (misura dei dati di diffrazione, soluzione del problema della fase, costruzione del modello e affinamento) sono stati resi molto più veloci rispetto al passato. Questo mette ancora più in evidenza il rallentamento introdotto dallo stadio di crescita dei cristalli. Naturalmente progressi sono stati compiuti anche in questo campo e le metodologie di cristallizzazione hanno beneficiato, per esempio, dell'introduzione di robot nella preparazione delle piastre di cristallizzazione usate nella tecnica di diffusione di vapore. Questi hanno anche permesso di ridurre significativamente le quantità di materiale usato nei test di cristallizzazione, riducendo al minimo la dimensione delle gocce con cui si eseguono le prove (50-100 nanolitri), il che permette di eseguire moltissimi test partendo da una limitata quantità di campione. In aggiunta, numerosi kit contenenti soluzioni precipitanti già pronte sono disponibili commercialmente, cosa che semplifica il lavoro del cristallografo. Tuttavia, l'ottenimento di cristalli di macromolecole rimane per ora, più che una scienza, un'arte nella quale vengono provate un grande numero di condizioni variando numerosi parametri: natura e concentrazione del precipitante, pH e tipo di tampone, temperatura, concentrazione di macromolecola, additivi e, eventualmente, detergenti. La corretta combinazione di questi fattori può portare all'ottenimento di cristalli adatti all'analisi ai raggi X. Se questo non avviene, la ragione spesso risiede nello stato del campione: per esempio, la macromolecola può contenere una porzione consistente di regioni flessibili o disordinate, o può tendere ad aggregare dando origine a oligomeri di dimensioni variabili. Infine, un discorso a parte meritano le proteine di membrana. Queste, che rappresentano una porzione consistente del genoma, intorno al 30%, sono difficili da produrre in grande quantità e ancora più difficili da cristallizzare.

In conclusione, lo stadio della cristallizzazione rappresenta oggi il collo di bottiglia nel processo della determinazione della struttura ai raggi X delle macromolecole. Questo potrebbe però essere superato, almeno parzialmente, non solo da possibili future innovazioni nelle tecniche di cristallizzazione, ma anche da un nuovo attore appena apparso sulla scena, il laser a raggi X a elettroni liberi (*X-ray Free-Electron Laser*, XFEL), discusso nel prossimo paragrafo.

### XFEL e nano-cristallografia

La nuova frontiera della cristallografia macromolecolare è rappresentata da una sorgente chiamata XFEL. È un acceleratore lineare di elettroni che produce impulsi coerenti di radiazione che durano qualche decina di femto-secondi e che sono caratterizzati da una brillantezza che supera di svariati ordini di grandezza quella degli attuali sincrotroni. Un XFEL caratterizzato da una lunghezza d'onda di circa 1 Å è già operativo nei pressi di Stanford (*Linear Coherent Light Source*, LCLS; Fig. 3) e altri sono in via di costruzione, tra i quali l'XFEL Europeo ad Amburgo (Germania).

L'XFEL produce brevi impulsi di raggi X così intensi che il cristallo di proteina viene totalmente distrutto in un tempo anche più breve dell'impulso medesimo; ciononostante, prima che il cristallo decada completamente si riesce ad acquisire uno spettro. Questo rappresenta solo una piccola parte dello spettro di diffrazione totale, poiché normalmente il cristallo deve essere ruotato per poter coprire l'intero spazio reciproco. Poiché l'interazione cristallo-radiazione ha luogo nel vuoto e i cristalli sono distrutti nel giro di pochi femto-secondi, è stata messa a punto una tecnica che consiste nell'iniettare micro- o nano-cristalli all'interno di uno spruzzo di liquido (*liquid jet*). Ogni volta che un impulso di raggi X incontra un cristallo, si osserva sul rivelatore uno spettro di diffrazione. Combinando i singoli spettri ottenuti da un grande numero di cristalli disposti con orientazioni casuali l'intero spettro tridimensionale può essere ricostruito e la struttura determinata.

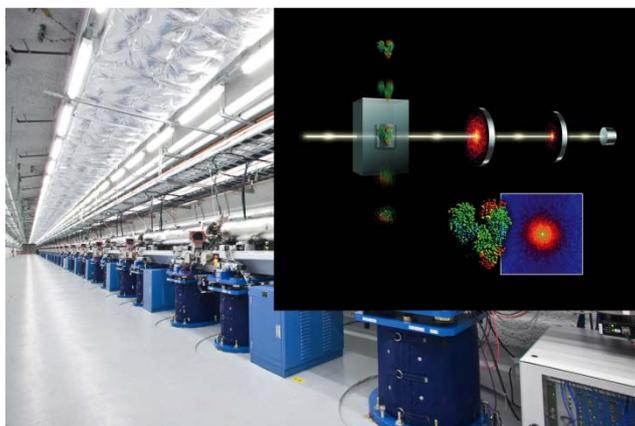


Fig. 3 - Vista di alcuni moduli del X-ray Free-electron laser SCLS operante a SLAC, Stanford University, USA (Cortesia SLAC National Accelerator Laboratory, Archives and History Office). Nell'insero, una rappresentazione idealizzata della procedura per ottenere immagini a risoluzione atomica di singole molecole

Il principale vantaggio di questa tecnica risiede nella possibilità di usare cristalli che, per le loro dimensioni, non fornirebbero uno spettro di diffrazione osservabile in un sincrotrone. Le prime prove sperimentali hanno mostrato che, per misurare un intero spettro, è necessario l'uso di più di un milione di *snap-shots*. Solo una piccola percentuale di questi, meno del 5%, incontrano in modo produttivo un cristallo dando origine a un'immagine di diffrazione, e solo una frazione di queste ultime può essere indicizzata; questo implica l'uso di centinaia di migliaia di singoli micro- o nano-cristalli. Non essendo possibile, almeno per ora, sincronizzare l'impulso dei raggi X con la presenza di un cristallo nel fascio medesimo, un grande numero di cristalli viene perso e un grande numero di frame del rivelatore risulta vuoto. Tuttavia, un paio di strutture sono state determinate, a dimostrazione della fattibilità del metodo<sup>7,8,9,10</sup>. Il vantaggio nell'usare micro- o nano-cristalli appare ovvio: la crescita di cristalli grandi e ben ordinati è il collo di bottiglia della biologia strutturale. In svariati casi, soprattutto nel caso di proteine di membrana, spesso non si ottengono mai cristalli adatti. Al contrario, il precipitato da una soluzione di proteina è generalmente non amorfo, ma consiste di un grandissimo numero di cristallini di dimensione nanometrica.

Infine, va aggiunta una considerazione teorica. La teoria cinematica della diffrazione dei raggi X assume che, per poter osservare spot di diffrazione in corrispondenza dei picchi di Bragg, il

cristallo sia infinito (e perfettamente ordinato). Questo naturalmente non è mai vero, ma in pratica un numero sufficientemente grande di “celle unitarie” è considerato un’approssimazione adeguata. Consideriamo come esempio un cristallo, di forma cubica, la cui dimensione per ogni lato sia di 1  $\mu$ . Se la cella cristallina ha dimensioni di 10 nm, questo dà una ripetizione di 100 celle unitarie in ciascuna delle tre direzioni; se lo stesso cristallo contiene celle di 100 nm, sono presenti solo 10 celle per ogni direzione. In entrambi i casi, le condizioni di Laue sono solo parzialmente soddisfatte e lo spettro apparirà come una sequenza di picchi di Bragg inframezzati da una serie di frange. Questo spettro di diffrazione continuo può rappresentare un vantaggio, poiché potrebbe essere usato per risolvere in modo diretto il problema della fase<sup>11</sup>.

### Cristallografia senza cristalli?

L’impossibilità di cristallizzare parecchie categorie di campioni, ad esempio macromolecole composte da domini connessi tra loro da regioni flessibili o complessi macromolecolari labili o transienti, impedisce di avere informazioni strutturali a livello atomico su campioni biologicamente molto importanti. XFEL offre l’opportunità, grazie alla coerenza della sua radiazione, di ottenere la struttura di singole molecole, senza la necessità che queste siano presenti all’interno di un cristallo. È, infatti, possibile ricostruire l’immagine di un oggetto dal suo spettro di diffrazione continuo usando appropriati algoritmi di recupero della fase<sup>12</sup>. Nel caso dell’esposizione di una singola molecola al fascio di un XFEL, nei pochi femtosecondi che precedono l’esplosione della molecola si raccoglie lo spettro, questa volta continuo (Fig. 3). Anche in questo caso, come per il cristallo, la radiazione diffusa raccolta sul rivelatore ci fornisce una proiezione bidimensionale della molecola, e numerose immagini di proiezioni diverse devono essere combinate per generare la struttura tridimensionale. La ricombinazione delle singole proiezioni è molto più complessa che nel caso dei cristalli, poiché in quest’ultimo caso l’indicizzazione dello spazio reciproco rende il processo semplice. Al momento, sono state effettuate solo simulazioni<sup>13</sup> e non conosciamo ancora il vero limite di risoluzione di questa tecnica di ricostruzione dell’immagine da particelle singole. Tuttavia, essa apre orizzonti ancora inesplorati alla biologia strutturale.

### Conclusioni

Dopo gli straordinari progressi tecnici degli ultimi decenni, la cristallografia macromolecolare ha raggiunto la maturità. Un settore un tempo coltivato da un numero limitato di specialisti è diventato una tecnica di uso comune tra i biologi molecolari e i biochimici che cercano di risolvere rilevanti problemi biologici. Ma i nuovi sviluppi tecnologici, quali l’*X-ray Free-Electron Laser*, lasciano intravedere che altre e ancor più straordinarie possibilità sono dietro l’angolo.

---

### Bibliografia

- <sup>1</sup> A.E.H. Tutton, *Nature*, 1922, **110**, 2749.
- <sup>2</sup> J.C. Kendrew *et al.*, *Nature*, 1958, **181**, 662.
- <sup>3</sup> M.F. Perutz *et al.*, *Nature*, 1960, **185**, 416.
- <sup>4</sup> C. Giacovazzo *et al.*, *Fundamentals of Crystallography*, Oxford University Press, Oxford, 2011.
- <sup>5</sup> M. Pioletti *et al.*, *EMBO J.*, 2001, **20**, 1829.
- <sup>6</sup> V.S. Reddy *et al.*, *Science*, 2011, **329**, 1071.
- <sup>7</sup> T.F. Lerch *et al.*, *Virology*, 2010, **26**, 403.
- <sup>8</sup> H.N. Chapman *et al.*, *Nature* 2011, **470**, 73.
- <sup>9</sup> S. Boutet *et al.*, *Science* 2012, **337**, 362.
- <sup>10</sup> L. Redecke *et al.*, *Science* 2013, **339**, 227.
- <sup>11</sup> R.A. Dilanian *et al.*, *Acta Cryst. Section A*, 2013, **A69**, 108.
- <sup>12</sup> J.R. Fienup, *Applied optics*, 1982, **21**, 2758.
- <sup>13</sup> A. Ravasio *et al.*, *Phys. Rev. Lett.*, 2009, **103**, 28104.